Research Memorandum
ETS RM–21-02

Exploring Math Education Relations by Analyzing Large Data Sets II

Jonathan Weeks
Patricia Baron

April 2021

ETS Research Memorandum Series

Exploring Math Education Relations by Analyzing Large Data Sets II

Jonathan Weeks and Patricia Baron
ETS, Princeton, New Jersey, United States

April 2021

Corresponding author: J. Weeks, E-mail: jweeks@ets.org

Suggested citation: Weeks, J., & Baron, P. (2021). *Exploring math education relations by analyzing large data sets II* (Research Memorandum No. RM-21-02). ETS.

Action Editor: Sooyeon Kim

Reviewers: Dan McCaffrey and Diego Zapata

Abstract

The current project, Exploring Math Education Relations by Analyzing Large Data Sets (EMERALDS) II, is an attempt to identify specific Common Core State Standards procedural, conceptual, and problem-solving competencies in earlier grades that best predict success in algebraic areas in later grades. The data for this study include two cohorts of California students with Smarter Balance (SB) mathematics scores. SB items were clustered based on a priori defined prealgebra knowledge domains and algebra content areas. Two approaches were then used to characterize student performance for several of these competencies: item residuals and factor scores. These performance indices were subsequently used in a series of regression models to examine the relationships between prealgebra competencies and later algebra outcomes. The findings from this study have implications for the use of the SB assessment for monitoring students' readiness for high school algebra and highlight areas that might be particularly important for success in this core mathematics domain. This report also provides direction for future studies.

Keywords: Common Core State Standards (CCSS), math, algebra, large data sets, Smarter Balanced, problem-solving competency

Acknowledgments

The authors would like to thank Dr. Mao Vang, director of the California Department of Education's (CDE's) Assessment Development and Administration division, and her staff, as well as Dr. Shanine Coats, director of the CDE's Curriculum Frameworks & Instructional Resources division, and her staff for their valuable input to earlier drafts of this report. We would also like to acknowledge the technical contributions by the EMERALDS Working Team and Dan McCaffrey; their suggestions during the interim analyses helped refine the methodology and guide the interpretation of the results.

Table of Contents

List of Tables

Page

List of Figures

Page

Overview

Competence with algebra is the foundation for learning the more complex mathematics demanded in science, technology, engineering, and mathematics (STEM) fields (National Mathematics Advisory Panel [NMAP], 2008) and now contributes to employability and wages in many blue-collar occupations (Bynner, 1997). Improving students' understanding of algebra has been a long-term educational priority in the United States; however, achieving this goal has been elusive (Stein et al., 2011), especially for students who have historically been underrepresented in STEM fields. For example, the eighth grade mathematics section of the 2019 National Assessment of Educational Progress (NAEP) defines basic skills as including conceptual and procedural competence with whole and rational numbers. Although these are critical mathematical competencies (Siegler & Braithwaite, 2017), it is unlikely that eighth graders performing at a NAEP basic level are prepared for a rigorous high school algebra course (NMAP, 2008). Yet, 69% of U.S. students performed at or below basic levels, with 86% and 80% of Black and Hispanic students, respectively, scoring at these levels. Overall, only 10% of eighth grade students have achieved the advanced competencies that well position them for rigorous high school mathematics coursework.

The content coverage of the mathematics NAEP assessment—a long-standing benchmark of U.S. students' educational progress—and more recent measures that incorporate the Common Core State Standards (CCSS) are not completely aligned in terms of content representation (e.g., more or less emphasis on fractions). However, overall patterns of performance are the same; most U.S. students are not fully prepared for a rigorous course in high school algebra (Daro et al., 2015; Hughes et al., 2019). The CCSS are based on mathematics standards from countries that consistently produce students who are well-educated in mathematics (e.g., NMAP, 2008; Schmidt & Houang, 2012) with a goal toward emphasizing critical procedures, concepts, and problem-solving skills that best prepare students for a demanding high school mathematics curriculum. By extension, the standards are expected to lay a foundation for college mathematics and entry into the workforce (Zimba, 2014). That said, while most mathematics content covered in elementary and middle school and highlighted in

CCSS will have utility in some contexts, it is possible that not all mathematics content is equally critical in terms of preparation for high school algebra.

The current project is an attempt to identify specific CCSS procedural, conceptual, and problem-solving competencies in earlier grades that provide the most critical foundation for success in algebraic areas in later grades. Gaining algebraic competence is undergirded by strong procedural and conceptual competencies in key areas, like fractions (Hurst & Cordes, 2018; Mou et al., 2016; Siegler et al., 2012). However, simple exposure to fractions and other aspects of arithmetic in earlier grades is insufficient to ensure adequate preparation for learning algebra. Rather, the quantity and quality of the opportunities to learn this content, in addition to a variety of other factors (e.g., family background, economic disadvantage, attentiveness in the classroom) should be taken into consideration (Bailey et al., 2014; Geary et al., 2017; Lee & Bull, 2016).

The available Smarter Balanced (SB) data for this project (described later) do not allow us to directly control for all these myriad influences; however, they do allow for isolating the relationship between particular competencies in elementary school and achievement in later grades while controlling for overall mathematics ability. As such, it may be feasible to identify skills that have a persistent effect in later grades and in future achievement. For this project, SB items were clustered based on a priori defined prealgebra knowledge domains and algebra content areas. Two approaches were then used to characterize student performance for several of these competencies: item residuals and factor scores. These performance indices were subsequently used in a series of regression models to examine the relationships between prealgebra competencies and later algebra outcomes.

The identification of key procedural, conceptual, and problem-solving competencies is a first step toward better preparing U.S. students for success in high school algebra. Depending on the relative importance of various skills in predicting later algebra performance, information about these competencies could be used to provide guidance in the development of mathematics frameworks, textbooks, and instructional foci in classroom settings, with a goal toward increasing performance across the competencies for all students.

Smarter Balanced Test

The SB computer adaptive test (CAT) algorithm was constructed using an underlying blueprint developed by mathematics content experts (for the mathematics section) and guided by the CCSS for mathematics. As noted previously, the goal of the latter was to provide a more focused and coherent mathematics education system that is comparable to that found in nations with consistently high-achieving students.

Among other things, the blueprint designates mathematics content as being major, supporting, or additional. These designations were intended to prioritize instructional time and the foci of educational assessments at each grade level. For instance, fifth grade students are expected to spend most of their time learning about place value, solving complex whole number and decimal arithmetic problems, and extending their conceptual understanding and operation skills with fractions. Supporting and additional instruction would include graphing quantitative relations in the coordinate plane and measurement.

The distribution of items on the SB assessment follows these instructional priorities, with about 75% of the elementary grade items focused on topics that are considered most critical in the progression toward algebra. The algebra-related items are designated as major content-area items in the blueprint, such as base-10 and fractions competencies in fifth grade. More broadly, these items largely assess fundamental, conceptual, and procedural knowledge within the content area but also include items that measure complex problem-solving, communicating and reasoning, and the ability to use modeling to solve real-world problems. The remaining items assess some of the additional and supporting clusters, but not all students receive a significant number of items in these areas due to the adaptive nature of the assessment.

The focus here is on major elementary grade content like fractions concepts that are thought to be foundational for later algebra learning and on major algebra topics covered in later grades. Competencies in other areas like measurement and geometry were also examined, allowing for an evaluation of the discriminant validity of the hypothesized major prealgebra competencies. For instance, discriminant validity would be demonstrated if fifth grade fractions competencies were a stronger predictor of later algebra outcomes than fifth grade geometry competencies.

Current Project

The core questions addressed by the Exploring Math Education Relations by Analyzing Large Data Sets (EMERALDS) II study are shown in Table 1.

Table 1. Core Research Questions for EMERALDS II

Letter	Question
A	Can we identify the factors in K–8 achievement data most predictive of success in algebra? Can we find factors more specific than “mathematics” and more useful than broad topics like “number” and “geometry”?
B	Are we spending too much time on some less important standards and not enough on some more important standards? Where should more time and effort be invested in mathematics instruction, and where less?
C	How do clusters of students (classified according to their profiles across assessment items) fare over time? Do the achievement gaps widen for some clusters (controlling for background factors) but not others? Students with different profiles may benefit differently from different interventions. Some topics (refer to Question A) may be more difficult for some profiles, while other topics may be more difficult for others. Are there some schools outperforming demographics for some clusters?
D	Can the factors (Question A), emphasis (Question B), student profiles (Question C), or trajectories in achievement differences among subpopulations be associated with the proportion of the variance in mathematics achievement among districts compared to among schools within districts, compared to among classrooms within schools, compared to among students within classrooms?

The primary focus of this study is on Question A, specifically, identifying the earlier mathematical competencies that are the most predictive of later algebra competencies, above and beyond the influence of general mathematics achievement. The associated analyses also allow for a preliminary assessment of Question C, that is, whether the overall patterns identified in Question A differ across various demographic subgroups. Note that the focus of Question C is not to simply rank order students in different demographic subgroups. Rather, the goal is to examine whether the same pattern of early relative strengths (e.g., early fractions knowledge), as related to later strengths in algebra (e.g., fluency of expression evaluation), is the same across subgroups. If this is the case, curricular changes that focus on these key early content and process (e.g., problem solving) areas may be relevant to better prepare all students for success in algebra. At the same time, identifying these key early areas may help identify whether certain student groups have deficits, relative to their overall achievement, in key areas and whether these deficits are systematically related to schools or districts. Such a finding would have implications for changes in curricular foci for such schools or districts.

Methodology

Data

The data for this study include two cohorts of California students with SB mathematics scores. The majority of these students also had SB English language arts (ELA) scores; however, this was not a requirement. The first cohort, hereafter referred to as the 5/8 cohort, includes 420,089 students with contiguous years of math scores in Grades 5, 6, 7, and 8 from 2016 to 2019 respectively. The second cohort, hereafter referred to as the 4/7 cohort, includes 429,968 students with contiguous years in Grades 4, 5, 6, and 7 from 2016 to 2019 respectively. The 5/8 cohort includes only students who had Grade 5 scores in Year 1, Grade 6 scores in Year 2, Grade 7 scores in Year 3, and Grade 8 scores in Year 4. The same holds for the corresponding grade levels in the 4/7 cohort. The cohorts do not include students who had missing scores, skipped a grade level, or were retained in the same grade during the interval; the demographic characteristics of the students excluded from the analyses are similar, proportionally, to the students included in the analyses. Tables 2 and 3 summarize the demographic makeup of these two cohorts, respectively.

Table 2. Student Counts by Gender and Demographic Group—5/8 Cohort

Demographic student group	Male	Female	Total
American Indian or Alaska Native	1,049	1,032	2,081
Asian	19,632	19,000	38,632
Native Hawaiian or Other Pacific Islander	923	973	1,896
Filipino	4,909	4,766	9,675
Hispanic or Latino	119,994	117,115	237,109
Black or African American	10,637	10,372	21,009
White	47,770	44,910	92,680
Two or more races	7,400	7,502	14,902
Ethnicity not reported	902	898	1,800
English only	116,277	111,908	228,185
Initial fluent English proficient	7,806	9,203	17,009
English learner	27,032	19,018	46,050
Reclassified fluent English proficient	62,090	66,420	128,510
English proficiency to be determined	8	9	17
English proficiency unknown	3	10	13
No special education services	182,222	190,592	372,814
Special education services	30,994	15,976	46,970
Not economically disadvantaged	82,364	79,886	162,250
Economically disadvantaged	130,852	126,682	257,534
Migrant	1,582	1,564	3,146
Not migrant	211,634	205,004	416,638
No available demographic information			305

Table 3. Student Counts by Gender and Demographic Group—4/7 Cohort

Demographic student group	Male	Female	Total
American Indian or Alaska Native	1,055	1,059	2,114
Asian	19,465	18,678	38,143
Native Hawaiian or Other Pacific Islander	953	921	1,874
Filipino	4,814	4,512	9,326
Hispanic or Latino	124,658	120,940	245,598
Black or African American	10,795	10,635	21,430
White	48,014	45,180	93,194
Two or more races	8,108	7,765	15,873
Ethnicity not reported	1,102	1,077	2,179
English only	121,774	115,864	23,7638
Initial fluent English proficient	7,796	9,296	17,092
English learner	32,325	24,365	56,690
Reclassified fluent English proficient	57,043	61,222	118,265
English proficiency to be determined	13	12	25
English proficiency unknown	13	8	21
No special education services	186,996	193,530	380,526
Special education services	31,968	17,237	49,205
Not economically disadvantaged	83,209	79,713	162,922
Economically disadvantaged	135,755	131,054	266,809
Migrant	1,558	1,523	3,081
Not migrant	217,406	209,244	426,650
No available demographic information			237

Item Clusters

An underlying assumption in the development of the SB math scale is that the measured construct is essentially unidimensional. That is, student performance within and across grade levels can be characterized by a single, dominant factor. However, it is feasible that some unique variability in student scores can be explained by performance on subsets of items that purport to measure various targeted skills. Content experts at Student Achievement Partners used information about the content standards, claims, and task models for items in the SB item pool to identify 14 item clusters (refer to Tables 4 through 8). For detailed information about the item clusters, refer to Tables 9 and 10.

To construct these item clusters, mathematics-content experts familiar with the SB assessment identified relevant groupings of items that varied by mathematical content and complexity. The first group of items was based on SB Claim 1, which focuses on concepts and procedures. These are the number sense predictors shown in Table 4; the complete set of claims and standards used to identify the items are shown in Appendix A. These items typically require students to execute mathematical procedures (e.g., solve multidigit multiplication problems), as well as explain and apply basic concepts (e.g., understanding place value in the base-10 system).

For several of the other item clusters, the experts focused on Claim 2 (problem solving) and Claim 4 (modeling and data analysis), which largely assess students' ability to solve basic and applied word problems, to take real-world problems and construct and use mathematical models to analyze these problems, and to reason and communicate about mathematics. Examples of these types of items are shown in Table 5. Within each of these broader categories (e.g., number sense), items were further differentiated based on arithmetical content and problem-solving complexity.

As a contrast to the set of algebra-related clusters, the content experts identified items that were not expected to be as strongly related to later algebra performance as the number sense, problem-solving, and reasoning items; these largely include geometry and measurement items (refer to Table 6). It was not possible to identify a set of items without some arithmetic operations; hence, this is not a pure contrast. In addition to this contrast cluster, the experts identified eight clusters that serve as potential Grade 4 and Grade 5 predictors of later Algebra I performance: Three clusters are associated with number sense skills and five associated with problem-solving and reasoning skills.

In addition to the potential predictors, four item clusters were identified that relate to critical components of Algebra I performance in Grades 7 and 8 (refer to Table 7). These are the primary outcome measures of interest in the analyses. An additional cluster associated with geometry, statistics, and probability was also identified as an outcome measure for the purpose of establishing a contrast (refer to Table 8).

The items identified for each of the predictor and outcome clusters were used to create cluster-specific subscores for use in a series of regression analyses. The approaches used to create the subscores and the regression analyses are presented later.

Table 4. Core Predictor of Algebra—Number Sense

Content area	Target	Example
A1a: Whole Numbers	Operations with and conceptual understanding of place value with whole numbers	Released Item
A1b: Fractions	Operations with and conceptual understanding of proper fractions and mixed numbers	Released Item
A1c: Decimals and Place Value	Read, write, and compare decimals to the thousands	Released Item

Table 5. Core Predictor of Algebra—Problem Solving

Content area	Target	Example
A2a: Basic problem solving: Whole numbers	One-step word problems involving whole numbers	Released item
A2b: Basic problem solving: Fractions	One-step word problems involving fractions	Released item
A3a: Complex problem solving: Whole numbers	Multistep or higher-complexity word problems with whole numbers	Released item
A3b: Complex problem solving: Rational numbers	Multistep or higher-complexity word problems with fractions or decimals	Released item
A4: Mathematical reasoning and communication	Construction of arguments to support mathematical reasoning or to critique the reasoning of others	Released item

Table 6. Core Contrast for Algebra—Geometry and Measurement

Content area	Target	Example
AG: Understand shapes, volume, and measurement	Convert like measurements within a measurement system; identify properties of shapes and volumes of solids	Released item

Table 7. Critical Components of Algebra I

Content area	Target	Example
B1: Quantitative literacy	Identify and graph relationships between quantities	Released item
B2: Algebra as generalized arithmetic	Read, write, and transform expressions and equations using arithmetic operations	Released item
B3: Algebra as functional thinking	Formulating and interpreting linear, quadratic, and exponential relations between quantities	Released item
B4: Algebra in constraint equations	Solve constraint equations involving linear, quadratic, and exponential relations between quantities	Released item

Table 8. Critical Contrast Outcome—Geometry, Statistics, and Probability

Content area	Target	Example
B5: Understand the properties of geometric figures and the basics of sampling, distributions, and inferences	Construct and describe the features of geometrical figures and the relations between them; use random sampling to draw inferences about a population	Geometry released item Statistics released item

Design

The data used in the analyses are based on combined data from the 5/8 cohort and the 4/7 cohort. Originally, the study design only included data from the 5/8 cohort; however, after a review of observed responses for each student, within each cluster, there were too few responses to defensibly estimate student-level indices (subscores) for certain clusters in Grade 5. Similarly, with the 4/7 cohort, there were several clusters with too few items to defensibly estimate student-level indices. The set of clusters with nonmissing indices partially overlapped between the 5/8 and 4/7 cohorts. As such, by including data from both cohorts, student-level indices for each cluster could be included in the regression models. The challenge then was to figure out how best to combine the data. The resolution was based, in part, on the approaches used to create the student-level indices.

Two approaches were used to characterize student performance within each cluster: an evaluation of item residuals and the creation of subscale scores. Item residuals are computed by taking the difference between the item responses and the expected probabilities of a correct response, or scoring in a particular category, for polytomously scored items. (A detailed description of the approach used to compute the residuals is presented in the next section, Item Residuals.) The expected probabilities were derived using the item parameters from the unidimensional item response theory (IRT) model for SB mathematics scale and the overall estimates of student mathematics ability, at the given grade level. These residuals were then averaged for each student, for each cluster, to create the indices used in the regression analyses. Note that it is possible to create average residuals for each cluster, with sufficient data, for students at each grade, within each cohort. This can be done for both the predictor and outcome clusters. However, this does not resolve the issue of how to include these variables in the regression models, particularly due to different proximal and distal grades across the two cohorts.

As an alternative to the residual approach, item parameters and student abilities were reestimated for each cluster. This supposes that the subscales can be treated as separate, unidimensional scales. In order to place the items and scores on the same scale, separately for each cluster, across grade levels and cohorts, it is necessary to impose identification constraints on the IRT model (refer to the IRT Subscales section for a description of the methodology). In short, these constraints fix the distribution of the item parameters or the scores (e.g., setting the ability distribution to be normal with a mean of 0 and standard deviation of 1). A full explication of linking designs and associated considerations is beyond the scope of this report, yet some background is needed to understand how the data were combined across cohorts. For the identification constraints to be meaningful, one of two conditions should be met: (a) a common set of items should be administered to students in different groups (i.e., grade levels within and between cohorts) or (b) the students should be part of the same underlying population. Given the adaptive design of the test and the differences between Grade 4 students from the 4/7 cohort and Grade 5 students from the 5/8 cohort, neither of these conditions holds. As such, it is not feasible to combine the data across cohorts—in order to obtain scores for all the identified clusters—without making additional assumptions.

To allow for the inclusion of scores across the identified clusters, a decision was made to combine the responses from Grades 4 and 5 from the 4/7 cohort with the data for Grade 5 from the 5/8 cohort. That is, responses from Grades 4 and 5 were treated as if they were collected as part of a single test administration. In the rare instances where a student received the same item in both years, the higher scored response was retained. These combined responses were added to the set of Grade 5 responses from the 5/8 cohort. The resulting data set included one row per student (a total of 850,057 rows) and columns for each unique item in each cluster. Similarly, responses from Grades 7 and 8 from the 5/8 cohort were combined with data for Grade 7 from the 4/7 cohort (refer to Figure 1). For the IRT analyses, the students in Grades 4 and 5 were assumed to be from the same underlying population; students in Grades 7 and 8 were also assumed to be from a single, separate population.

Figure 1. Design for Combining Data From the 5/8 and 4/7 Cohorts

5/8 Cohort		5		7	8
4/7 Cohort	4	5		7	

It is important to note that the defensibility of concatenating the item responses for the Grades 4 and 5 students and the Grades 7 and 8 students is premised on the assumptions that the same construct is measured at both grades and that student performance across grades is highly correlated. The scores in Grades 4 and 5 from the 4/7 cohort are correlated at $r = .86$; the scores in Grades 7 and 8 from the 5/8 cohort are correlated at $r = .87$. Although these correlations are fairly high, the shared variance between grades is only around 75%. It is unclear how much of the unexplained variance is due to random error, systematic differences associated with student learning, and/or potential changes in the construct. As such, concatenating the responses may be reasonable to allow for the inclusion of a more complete set of subscores in the regression analyses; however, any difference across the cohorts could affect the prediction of algebra scores.

The design for the residual analyses and the creation of subscores relied on the combined data across the cohorts. The residuals were computed using the ability estimates for each student in the respective grade and cohort, but the average residuals were aggregated across all available responses per cluster. As such, the residuals for the combined Grades 4 and 5 responses assume more of a composite underlying ability. On the other hand, the IRT subscales implicitly assume a single, composite ability that gives rise to the observed responses within each cluster. For this reason, when considering the results, the predictors should be thought of as a combined set of competencies in Grades 4 and 5 whereas the outcomes should be thought of as competencies in Grades 7 and 8.

Tables 9 and 10 summarize the number of items associated with each cluster in the combined data set, the number of students with at least three responses, the range in the number of observed responses, the mean number of responses, and the marginal score reliability. For example, for cluster A1b (Number Sense: Fractions), there were 117 items with observed responses; 108 were dichotomously scored, and nine were polytomously scored. Of the 850,057 students in the combined data set, 632,374 had three or more responses. Across all 850,057 students, the number of responses ranged from 0 to 9, with an overall mean of 3.51 responses. After excluding students with responses to fewer than three items, the mean number of responses increased to 4.19. Hence, even though the number of identified items for the cluster is quite large, due to the SB CAT design, each student received relatively few of these items. Despite the small number of observed responses per student for these students, the marginal reliability is 0.61.

Table 9. Cluster Information for Predictor Variables

Variable	A1a	A1b	A1c	A2a	A2b	A3a	A3b	A4	A_G
Number of items	74	117	113	3	86	161	113	379	142
Dichotomous	74	108	113	3	82	133	63	287	141
Polytomous	0	9	0	0	4	28	50	92	1
N students (w/3+ responses)	27,378	632,374	246,416	0	158,922	473,373	604,023	849,875	714,222
Min num responses	0	0	0	0	0	0	0	0	0
Max num responses	4	9	4	2	8	15	11	20	11
Mean num responses	0.90	3.51	2.26	0.02	1.38	3.94	3.54	14.68	4.28
Mean num responses (3+)	3.01	4.19	3.00	NA	3.35	6.43	4.25	14.69	4.77
Reliability	0.42	0.61	0.30	NA	0.46	0.64	0.55	0.74	0.48

Table 10. Cluster Information for Outcome Variables

Variable	B	B1	B2	B3	B4	B5
Number of items	521	283	118	114	6	74
Dichotomous	498	262	118	112	6	73
Polytomous	23	21	0	2	0	1
N students (w/3+ responses)	849,775	849,733	727,779	617,096	0	431,431
Min num responses	0	0	0	0	0	0
Max num responses	29	22	5	10	2	5
Mean num responses	17.15	10.84	2.85	3.41	0.04	2.62
Mean num responses (3+)	17.15	10.85	3.00	4.09	NA	3.24
Reliability	0.86	0.80	0.48	0.33	NA	0.25

With respect to the marginal reliabilities, the presumption is that each student has responses for the same number of items. Although this is generally true, there are some clusters with notably greater variability in the number of responses. As such, the reported marginal reliabilities should be interpreted as rough indicators of stability. Note that the outcome measure B is based on all the items from clusters B1, B2, B3, and B4; this serves as an overall measure of algebra performance. Additionally, all the predictor and outcome clusters were included in the regression analyses, with the exception of A2a and B4.

Item Residuals

As described earlier, the SB assessment uses an adaptive testing procedure. The measure provides an estimate of students' overall grade-level mathematical ability based on performance on a fixed number of items; however, the specific items vary from person to person. The nature of the assessment thus complicates the determination of student strengths and weaknesses for particular subskills. Stated differently, each student has relatively few item responses for a given skill. This makes it difficult to examine the psychometric properties of the collection of items for each competency directly. One potential solution is to use students' deviation from expected performance on each item (based on overall grade-level mathematics scores) to determine their relative strengths in core prealgebra areas and in later core algebra outcomes for the SB item standards identified by content experts.

Calculating and compiling item residuals requires a multistep approach. As a first step, SB item parameters and overall estimates of student mathematics ability were used to compute the expected probability of a correct response (or the probability of scoring in a particular category). The probability of a correct response for selected-response items (e.g., multiple-choice, true or false) was modeled using the two-parameter logistic (2PL) model (Birnbaum, 1968).

The 2PL model is given by

$$P_i(\theta_j) = exp[Da_i(\theta_j - b_i)] / (1 + exp[Da_i(\theta_j - b_i)]), \quad (1)$$

where $P_i(\theta_j)$ is the probability of a correct response to item *i* by a student with an overall mathematics competence θ_j, and a_i is the item discrimination parameter and b_i is the item difficulty parameter. The student (θ_j) parameters were provided as part of the data for each student and each academic year. The item parameters (a_i and b_i) were provided by SB for each item. *D* is a constant that puts the item parameters and ability estimates on the same metric as a normal ogive model (D = 1.7).

For constructed-response items—items scored in categories ranging from 0 to 5—the generalized partial credit model (GPCM; Muraki, 1992) was used to estimate item parameters. The GPCM is given by

$$P_{ih}(\theta j) = \frac{exp \sum_{v=1}^{h}[Da_i(\theta_j - b_i + d_{iv})]}{\sum_{h=1}^{n_i} exp \sum_{v=1}^{h}[Da_i(\theta_j - b_i + d_{iv})]}, \quad (2)$$

where $P_{ih}(\theta_j)$ is the probability of student *j* with overall mathematics ability θ_j obtaining a score of *h* on item *i*, *n* is the number of item categories, b_i is the item location parameter, d_{iv} is the category parameter for item *i* for category *v*, and *D* is the scaling constant given previously.

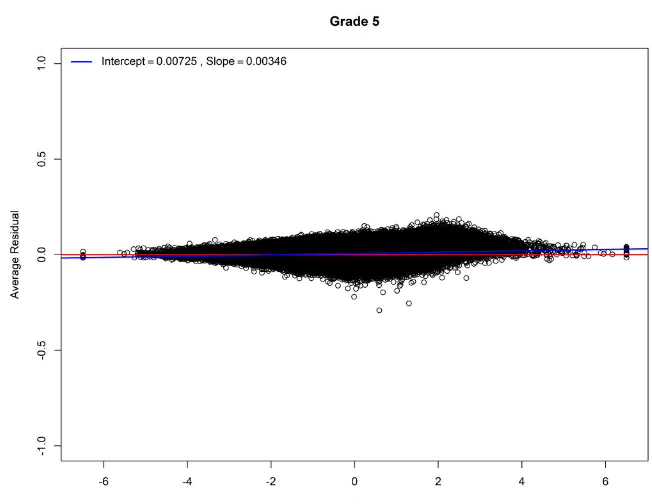
Response probabilities were computed for all items in the data set. To calculate the deviation between the student's score and the expected probabilities, we subtracted the probability from the scored response then divided the resulting score by the number of categories, minus one. These deviations are referred to as item residuals. The residuals are all on the same scale and vary between -1 and 1. Note that positive residuals are observed for correct responses (i.e., a correct response scored as 1 will always be greater than the expected probability of a correct response). Conversely, negative residuals are observed for incorrect responses. For polytomous items, scored responses in higher categories will generally correspond to higher residuals and vice versa.

As a next step, the residuals were averaged across the items associated with each item cluster for the predictor and outcome variables, respectively. These average residuals were standardized to have a mean of 0 and standard deviation of 1. The goal of the residual analysis is to create indices that capture the unique subskills that are distinct from general math ability. Stated differently, the residuals are meant to separate the variability in student skills that is not captured by the variability in overall ability. Figures 2 and 3 show the relationship between overall math ability, θ, and the average residuals for Grade 5 and Grade 8 students, respectively, from the 5/8 cohort. These representations are consistent for the other grades and the 4/7 cohort. The average residuals are centered around 0 and, on the whole, do not vary systematically, conditional on ability level.

The residual approach is premised on the notion that student performance for a given cluster can be meaningfully separated from overall performance. To the extent that the construct is essentially unidimensional, average item residuals should be smaller. On the other hand, if there is substantive variability in the clusters that is not explained by the overall score, one should expect to see larger average residuals for those clusters. If the average residuals by cluster are small, identifying relationships between proximal clusters (predictor clusters in the earlier grades) and distal clusters (outcome clusters in the later grades) may not be very informative. Further, because the same item scores are used for the clusters and the total scores, the residual approach can underestimate the relationship between early skills and later achievement in algebra.

As an alternative to the residual approach, item responses could be modeled separately for each cluster. If the measure is essentially unidimensional, this does not fully resolve the problem because the cluster subscores would still be highly correlated. Similar to the use of residuals is whether there is enough unexplained variability in the subscores in the earlier grades to provide meaningful inferences about expected algebra performance in later grades. With the residual approach, the goal is to partial out the variability associated with a subset of items from the overall variability in the data. Conversely, by only including a particular subset of items when creating a scale, only the associated variability in those data is considered. Although this is a subtle difference, the approach may provide more interpretable results with respect to examinations of subscore reliability and the practical significance of any associated effects in the regression analyses. However, the cluster-specific subscores do not separate out the overall ability from the specific skills of the item cluster. Consequently, the models for predicting outcomes must be adjusted to control for overall ability. Again, the results will be biased to some unknown extent because the cluster-specific subscores and the total ability measures will rely on overlapping items. Through a series of models, we can provide a more thorough exploration that ameliorates, in part, the sensitivity to potential biases.

Figure 2. Average Residuals by Ability for Grade Five Students From the 5/8 Cohort

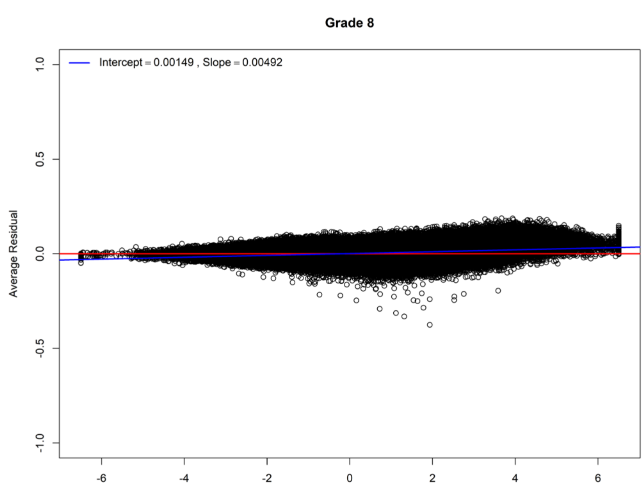


IRT Subscales

In the computation of item residuals, SB item parameters from the unidimensional IRT scale were used. These same item parameters could be used to create a set of separate subscores for each cluster; however, this approach does allow for deviations from the scale that might be explained by unique variability associated with the particular aspect of the construct measured by each item cluster. Item parameters and student abilities were estimated separately for each cluster (i.e., separate unidimensional scales were created); a multidimensional model was not employed. In general, all available data were used to estimate the item and person parameters; however, items with fewer than 100 responses and students with fewer than three responses, for a given cluster, were excluded from the estimation.

Dichotomously scored items were fit using the 2PL model; polytomously scored items were fit using the GPCM. A scaling constant of 1.7 was used to place the estimates on a normal metric. Item parameters were estimated via marginal maximum likelihood using the program mdltm (von Davier, 2017) based on a single group design. That is, the combined data across cohorts were treated as a single population. For the purpose of identification, the item slopes were constrained to have a mean of 1; the item difficulty/location parameters were constrained to have a mean of 0. Expected a posteriori estimates of student ability (factor scores) were compiled and standardized to have a mean of 0 and standard deviation of 1. Convergence of the estimation algorithm was assured before the resulting parameter estimates were used in further analyses. The end result of these analyses was the creation of 13 separate scales (scales were not created for A2a and B4 due to too few responses per student).

Figure 3. Average Residuals by Ability for Grade 8 Students From the 5/8 Cohort



Regression Analyses

By combining the data across cohorts, scores were available for a large number of students for all but two of the predictor and outcome clusters. To maximize the information available in the regression analyses, a multiple imputation approach was used to fill in the missing scores for the predictor variables (for both the residual indices and the cluster-specific factor scores). The R package Amelia (Honaker et al., 2011) was used for the imputation. Five imputed data sets were created. Tables 11 and 12 show the pairwise correlation between the predictor and outcome variables for the factor scores, including overall mathematics (math) and English language arts (ELA) scores, before and after the imputation respectively. The median absolute difference between the correlations is 0; the root mean squared difference is 0.008. In short, there are essentially no differences. Similarly, there were essentially no differences between the correlations for the residual indices.

Five separate regression models were fit for each of the four outcome measures, in addition to an overall measure of Algebra I performance, Cluster B, which includes all available items from B1, B2, B3, and B4. The model specification for each of the models was the same for each outcome measure, for both the residuals and the factor scores (i.e., the only difference for a given model is the outcome variable). An overview of the five models follows.

Model 1 only includes the set of subscores (residuals or factor scores) as predictors. The magnitude of the effect for each cluster will be largest (in absolute value) with this model. If the effects for specific clusters are not statistically nor practically significant under this model, we would not expect them to be meaningful predictors in any of the other models. Models 2–4 build on this first model by controlling for overall ability in Grade 5. As noted previously, one issue with the inclusion of overall math scores as a predictor, particularly in the models that use factor scores, is that the items used to create the subscores are the same items used in the overall math score. Models 5A–5D are intended to provide information about the effects for the meaningful clusters for various demographic subgroups. The first four models primarily address Question A, whereas Models 5A–5D speak more to Question C, as follows:

- Model 1—Main effects only for each of the predictor variables
    - A1a, A1b, A1c, A2b, A3a, A3b, A4, AG
- Model 2—Main effects only for each of the predictor variables with ELA score as a covariate
    - A1a, A1b, A1c, A2b, A3a, A3b, A4, AG, ELA (Grade 5)
- Model 3—Main effects only for each of the predictor variables with overall math score as a covariate
    - A1a, A1b, A1c, A2b, A3a, A3b, A4, AG, math (Grade 5)
- Model 4—Main effects only for each of the predictor variables with ELA and math scores as covariates
    - A1a, A1b, A1c, A2b, A3a, A3b, A4, AG, ELA (Grade 5), math (Grade 5)
- Model 5A—Main effects for A4 (the strongest predictor of algebra, based on the previous models), Grade 5 ELA score, and a range of demographic variables; additionally, interaction effects between A4 and the demographic variables (the interaction effects are denoted by the colons, e.g., A4: female is the effect for females with a given A4 score)
    - A4, female, Hispanic, or Latino, American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islander, Filipino, limited English proficient (LEP), 504 plan, economically disadvantaged, ELA (Grade 5), A4: female, A4: Hispanic or Latino, A4: American Indian or Alaska Native, A4: Asian, A4: Black or African American, A4: Native Hawaiian or Other Pacific Islander, A4: Filipino, A4: LEP, A4: 504 plan, A4: economically disadvantaged
- Models 5B, 5C, and 5D—Duplicate Model 5A, but use A3b, A1b, and A3a as the primary predictor, respectively

Table 11. Pairwise Correlations Based on Nonimputed Scores—Factor Scores

Factor	Math	ELA	A1a	A1b	A1c	A2b	A3a	A3b	A4	AG	B	B1	B2	B3	B5
Math	1.00	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ELA	0.81	1.00	-	-	-	-	-	-	-	-	-	-	-	-	-
A1a	0.44	0.39	1.00	-	-	-	-	-	-	-	-	-	-	-	-
A1b	0.65	0.51	0.28	1.00	-	-	-	-	-	-	-	-	-	-	-
A1c	0.26	0.18	0.15	0.07	1.00	-	-	-	-	-	-	-	-	-	-
A2b	0.40	0.32	0.19	0.16	0.10	1.00	-	-	-	-	-	-	-	-	-
A3a	0.65	0.59	0.38	0.41	0.16	0.29	1.00	-	-	-	-	-	-	-	-
A3b	0.72	0.62	0.33	0.43	0.12	0.26	0.49	1.00	-	-	-	-	-	-	-
A4	0.85	0.74	0.43	0.53	0.20	0.34	0.62	0.62	1.00	-	-	-	-	-	-
AG	0.34	0.30	0.17	0.10	0.10	0.11	0.24	0.18	0.27	1.00	-	-	-	-	-
B	0.80	0.72	0.40	0.52	0.21	0.33	0.56	0.61	0.73	0.28	1.00	-	-	-	-
B1	0.76	0.67	0.38	0.50	0.19	0.32	0.53	0.58	0.70	0.26	0.96	1.00	-	-	-
B2	0.42	0.39	0.22	0.25	0.12	0.17	0.29	0.31	0.38	0.17	0.58	0.42	1.00	-	-
B3	0.41	0.39	0.19	0.26	0.10	0.16	0.28	0.31	0.37	0.15	0.53	0.39	0.25	1.00	-
B5	0.40	0.38	0.19	0.26	0.09	0.18	0.28	0.31	0.37	0.14	0.39	0.37	0.20	0.21	1.00

Note. ELA = English language arts.

Table 12. Mean Pairwise Correlation Based on Imputed Missing Values (Five Imputations)—Factor Scores

Factor	Math	ELA	A1a	A1b	A1c	A2b	A3a	A3b	A4	AG	B	B1	B2	B3	B5
Math	1.00	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ELA	0.81	1.00	-	-	-	-	-	-	-	-	-	-	-	-	-
A1a	0.42	0.37	1.00	-	-	-	-	-	-	-	-	-	-	-	-
A1b	0.66	0.51	0.26	1.00	-	-	-	-	-	-	-	-	-	-	-
A1c	0.26	0.18	0.12	0.07	1.00	-	-	-	-	-	-	-	-	-	-
A2b	0.39	0.30	0.15	0.15	0.10	1.00	-	-	-	-	-	-	-	-	-
A3a	0.64	0.58	0.33	0.41	0.15	0.26	1.00	-	-	-	-	-	-	-	-
A3b	0.72	0.62	0.31	0.43	0.12	0.25	0.49	1.00	-	-	-	-	-	-	-
A4	0.85	0.74	0.41	0.53	0.20	0.33	0.62	0.62	1.00	-	-	-	-	-	-
AG	0.34	0.30	0.14	0.10	0.10	0.11	0.23	0.18	0.27	1.00	-	-	-	-	-
B	0.80	0.72	0.38	0.52	0.21	0.32	0.56	0.61	0.73	0.28	1.00	-	-	-	-
B1	0.76	0.67	0.36	0.51	0.19	0.30	0.53	0.58	0.70	0.26	0.96	1.00	-	-	-
B2	0.42	0.39	0.20	0.25	0.12	0.16	0.29	0.31	0.38	0.17	0.58	0.42	1.00	-	-
B3	0.41	0.39	0.19	0.26	0.10	0.16	0.29	0.31	0.38	0.15	0.53	0.40	0.25	1.00	-
B5	0.40	0.38	0.19	0.27	0.09	0.16	0.28	0.31	0.37	0.14	0.39	0.37	0.20	0.21	1.00

Note. ELA = English language arts.

Results

Item Clusters

To summarize the information in Tables 9 and 10, the number of items in the predictor clusters ranges from 74 to 379, while the number of responses, per student, ranges from 0 to 15, with an overall average number of responses (for students with a minimum of three responses) of 6.88. The number of items in the outcome clusters ranges from 74 to 521; the number of responses, per student, ranges from 0 to 22, with an overall average number of responses (for students with a minimum of three responses) of 5.84. The marginal score reliability for the A4 factor score is 0.74; the reliabilities for the factor scores for A1b, A3a, and A3b are around 0.6. The reliabilities for the other predictors range from 0.30 to 0.48. The reliability of the overall Algebra I outcome cluster is 0.86. The reliability of B1 is 0.8, while the reliabilities for the other outcome clusters range from 0.25 to 0.48. Note that there is no established method for estimating the reliability of residuals in this context.

Residuals

Table 13 shows the mean pairwise correlations for the residual indices (based on the imputed data). Overall, the correlations are quite small, with most values falling below 0.1. There are some correlations like {A1a, A3a}, {A1a, A4}, and {A3a, A4} that suggest some dependency. Note that this likely has more to do with the imposed orthogonality (removing the correlations between the clusters) than the information contained within in the residuals. Stated differently, it is unclear whether the low correlations are best explained by the imposed orthogonality or an actual lack of unique variance. Regression analyses for each of the models were still conducted (refer to Appendix B), although the explained variance was essentially 0. This is not surprising given that the predictor and outcome indices are basically uncorrelated. For this reason, the factor scores are used as the primary indices for reporting the regression results and evaluating the research questions.

Table 13. Mean Pairwise Correlation Based on Imputed Missing Values (Five Imputations)—Residuals

Factor	A1a	A1b	A1c	A2b	A3a	A3b	A4	AG	B	B1	B2	B3	B5
A1a	1.000	-	-	-	-	-	-	-	-	-	-	-	-
A1b	0.086	1.000	-	-	-	-	-	-	-	-	-	-	-
A1c	0.046	-0.033	1.000	-	-	-	-	-	-	-	-	-	-
A2b	0.008	-0.012	-0.025	1.000	-	-	-	-	-	-	-	-	-
A3a	0.162	0.034	-0.031	-0.006	1.000	-	-	-	-	-	-	-	-
A3b	0.022	-0.032	-0.057	-0.025	0.038	1.000	-	-	-	-	-	-	-
A4	0.133	-0.027	-0.049	-0.030	0.113	-0.022	1.000	-	-	-	-	-	-
AG	0.043	-0.054	-0.049	-0.033	0.040	-0.037	0.012	1.000	-	-	-	-	-
B	0.000	0.001	-0.001	0.000	0.001	0.000	0.000	0.000	1.000	-	-	-	-
B1	0.001	0.001	0.000	0.001	0.002	-0.001	0.001	-0.001	0.819	1.000	-	-	-
B2	0.000	0.001	-0.002	-0.001	0.000	0.002	0.000	0.002	0.492	0.118	1.000	-	-
B3	-0.003	0.000	-0.001	0.002	0.000	0.000	-0.001	0.001	0.464	0.109	0.059	1.000	-
B5	-0.001	-0.002	0.001	-0.001	0.001	0.000	-0.001	0.000	0.112	0.102	0.027	0.059	1.000

Regression Analyses

The standardized coefficients and corresponding standard errors for the factor scores (across the imputed data sets) for each of the five models are presented in Appendix C. As noted previously, the results for the regression analyses focus exclusively on the models using the factor scores. The overall percentage of variance in the outcome measures explained by the predictor models is shown in Table 14. Note, when these values are adjusted by the corresponding reliability estimates, the explained variances range from overall (.66–.77); quantitative literacy (.64–.75); generalized arithmetic (.34–.38); functional thinking (.49–.55); geometry and statistics (.62–.68); and ELA (.65–.83). After accounting for score reliability, the explained variance for functional thinking and geometry and statistics is notably higher. This suggests that if the subscores for these clusters were more reliable (i.e., students took more items within these clusters), their relationship to later algebra outcomes might be more meaningful. The standardized coefficients for Models 1 through 4 are shown in Tables 15 through 18, respectively. All coefficients that are not statistically significant at the $p \leq 0.05$ level at minimum are noted. Standard errors are in Appendices B and C. These are cluster adjusted standard errors, using district as the clustering variable (Bell & McCaffrey, 2002).

There are six outcome measures for each of these models. They include the four substantive algebra skills and the contrast measure (geometry and statistics). As an additional contrast, ELA scores from Grade 7 or 8 are considered, for each cohort respectively. Model 1 includes subscores for each of the predictor variables as main effects, Model 2 builds on Model 1 by including ELA scores from Grade 5 as a covariate. Model 3 is a variant of Model 2 that uses overall Grade 5 math scores as a covariate. Model 4 includes both ELA and math scores as covariates. The results for Models 5A through 5D are presented in Appendix C. Since Model 1 does not control for overall math or ELA ability, this model provides the largest effect sizes for the relationships between early prealgebra competences and later algebra outcomes. The contrast for the prediction of later algebra outcomes and later geometry and statistics and ELA outcomes (refer to Table 19) provides information on the specificity of the relationships between early prealgebra competencies and later algebra performance.

Table 14. Regression Models *R*-Squared Statistics by Outcome Variable—Factor Scores

Regression model	Overall algebra	Quantitative literacy	Generalized arithmetic	Functional thinking	Geometry & statistics	ELA
1	0.618	0.564	0.164	0.162	0.156	0.522
2	0.641	0.581	0.176	0.177	0.168	0.659
3	0.656	0.597	0.179	0.175	0.164	0.551
4	0.665	0.603	0.184	0.183	0.171	0.662
5A	0.617	0.558	0.175	0.174	0.169	-
5B	0.581	0.523	0.164	0.167	0.163	-
5C	0.573	0.515	0.161	0.165	0.161	-
5D	0.572	0.513	0.163	0.165	0.161	-

Note. ELA = English language arts.

Table 15. Model 1 Standardized Regression Coefficients by Outcome Variable—Factor Scores

Predictor variables	Overall algebra beta	Quantitative literacy beta	Generalized arithmetic beta	Functional thinking beta	Geometry & statistics beta
A1a—Whole numbers	0.059	0.056	0.037	0.023	0.025
A1b—Fractions	0.159	0.158	0.058	0.074	0.084
A1c—Decimals	0.057	0.053	0.038	0.027	0.021
A2b—PS fractions	0.071	0.069	0.033	0.028	0.042
A3a—CPS whole numbers	0.095	0.087	0.049	0.046	0.049
A3b—CPS fractions	0.198	0.192	0.093	0.103	0.103
A4—Reasoning & communicating	0.386	0.367	0.208	0.207	0.188
AG—Geometry & measurement	0.080	0.069	0.068	0.051	0.042

Note. CPS = complex problem solving; PS = problem solving.

Table 16. Model 2 Standardized Regression Coefficients by Outcome Variable—Factor Scores

Predictor variables	Overall algebra beta	Quantitative literacy beta	Generalized arithmetic beta	Functional thinking beta	Geometry & statistics beta
A1a—Whole numbers	0.050	0.048	0.030	0.016	0.018
A1b—Fractions	0.127	0.131	0.035	0.048	0.061
A1c—Decimals	0.050	0.046	0.032	0.021	0.015
A2b—PS fractions	0.059	0.060	0.024	0.019	0.034
A3a—CPS whole numbers	0.062	0.059	0.026	0.020	0.025
A3b—CPS fractions	0.146	0.149	0.056	0.061	0.066
A4—Reasoning & communicating	0.283	0.281	0.134	0.124	0.114
AG—Geometry & measurement	0.056	0.049	0.050	0.032	0.025
ELA	0.252	0.213	0.181	0.204	0.182

Note. CPS = complex problem solving; ELA = English language arts; PS = problem solving.

Table 17. Model 3 Standardized Regression Coefficients by Outcome Variable—Factor Scores

Predictor variables	Overall algebra beta	Quantitative literacy beta	Generalized arithmetic beta	Functional thinking beta	Geometry & statistics beta
A1a—Whole numbers	0.040	0.038	0.025	0.012	0.016
A1b—Fractions	0.020	0.030	-0.029	-0.006	0.021
A1c—Decimals	0.010	0.009	0.008	-0.001 [a]	-0.001[a]
A2b—PS fractions	0.015	0.018	-0.003 [a]	-0.004	0.017
A3a—CPS Whole numbers	0.049	0.045	0.020	0.020	0.028
A3b—CPS fractions	0.064	0.069	0.009	0.026	0.043
A4—Reasoning & communicating	0.160	0.159	0.066	0.077	0.086
AG—Geometry & measurement	0.016	0.010	0.027	0.014	0.013
Math	0.544	0.502	0.340	0.313	0.245

Note. CPS = complex problem solving; PS = problem solving.
[a] Not statistically significant.

Table 18. Model 4 Standardized Regression Coefficients by Outcome Variable—Factor Scores

Predictor variables	Overall algebra beta	Quantitative literacy beta	Generalized arithmetic beta	Functional thinking beta	Geometry & statistics beta
A1a—Whole numbers	0.037	0.036	0.022	0.009	0.013
A1b—Fractions	0.023	0.032	-0.027	-0.004	0.023
A1c—Decimals	0.013	0.011	0.010	0.002	0.002 [a]
A2b—PS fractions	0.017	0.019	-0.001 [a]	-0.002 [a]	0.019
A3a—CPS whole numbers	0.035	0.034	0.009	0.006	0.015
A3b—CPS fractions	0.053	0.060	0.000 [a]	0.015	0.033
A4—Reasoning & communicating	0.130	0.135	0.043	0.048	0.059
AG—Geometry & measurement	0.010	0.006	0.023	0.009	0.008
ELA	0.165	0.130	0.129	0.160	0.151
Math	0.454	0.430	0.269	0.225	0.163

Note. CPS = complex problem solving; ELA = English language arts; PS = problem solving.
[a] Not statistically significant.

Table 19. Models 1–4, Standardized Regression Coefficients and Standard Errors With ELA as the Outcome Variable—Factor Scores

Predictor variables	Model 1 beta	Model 2 beta	Model 3 beta	Model 4 beta
A1a—Whole numbers	0.041	0.018	0.024	0.013
A1b—Fractions	0.108	0.031	-0.012	-0.004
A1c—Decimals	0.029	0.010	-0.012	-0.002
A2b—PS fractions	0.043	0.016	-0.005	0.001
A3a—CPS whole numbers	0.120	0.040	0.080	0.031
A3b—CPS fractions	0.185	0.061	0.070	0.029
A4—Reasoning & communicating	0.371	0.124	0.175	0.072
AG—Geometry & measurement	0.086	0.027	0.030	0.012
ELA	-	0.608	-	0.579
Math	-	-	0.471	0.154

Note. CPS = complex problem solving; ELA = English language arts; PS = problem solving.

As shown in Table 14, all the models were strong predictors of later B1—overall algebra and B2—quantitative literacy. The results for these two outcomes are similar due to the high correlation between them (r = .96; refer to Table 12). The explained variance in the prediction of generalized arithmetic and functional thinking as well as geometry and statistics is considerably lower, but the predictors are still statistically significant.

The results from Models 1 and 2 suggest that A4—reasoning & communicating is the strongest predictor of algebra performance across the outcome measures. This is followed by A3b—CPS fractions, A3a—CPS whole numbers, and A1b—fractions. Note that these four clusters have the highest marginal reliability. Before, controlling for overall ability, the coefficient for A4—reasoning & communicating on overall algebra performance is 0.386 (a moderate effect size). The coefficients for A3b—CPS fractions, A3a—CPS whole numbers, and A1b—fractions are 0.198, 0.159, and 0.095 respectively (small effect sizes).

After controlling for ELA performance, the order of these three predictors remains unchanged; however, the magnitude of the coefficients is notably reduced—0.283, 0.146, 0.127, and 0.062, respectively. These are all small effect sizes, and the coefficient for A3a—CPS whole numbers becomes negligible. The coefficients are even smaller when overall math scores are included as a covariate—0.160, 0.064, 0.020, and 0.049, respectively. These coefficients are very small. This is likely due, in part, to lower subscore reliability (as discussed previously) and contamination; the overall math score includes performance on the items from each of the clusters. Note also that the ordering of the effect sizes for A3a—CPS whole numbers and A1b—fractions changes.

Based on the findings from Models 1 and 2, A4—reasoning & communicating, A3b—CPS fractions, A3a—CPS whole numbers, and A1b—fractions were included in Models 5A, 5B, 5C, and 5D, respectively, (refer to the tables in Appendix C) as primary predictors of algebra performance with ELA scores and demographic variables as additional covariates. The coefficients reveal notable positive effects for Asian students and negative effects for other minority students, LEP students, and economically disadvantaged students. For the latter groups, this may be due to the higher reading load for the more complex items.

Given the high correlation between overall algebra and quantitative literacy and the lower overall variance explained in the prediction of generalized arithmetic and functional thinking, we focused on the contrast between overall algebra and geometry and statistics and later ELA scores. The estimates for the prealgebra prediction of overall algebra are about double those for the prediction of geometry and statistics, suggesting earlier competencies are better predictors of later algebra than geometry and statistics. That said, if one accounts for the lower score reliability for geometry and statistics, the magnitude of the coefficients would be similar to overall algebra and quantitative literacy. This reduces the utility of geometry and statistics as a contrast measure. As such, there is little evidence that any of these predictors is uniquely important nor that later algebra performance is more substantively related to the identified content of the prealgebra measures.

Most of the prealgebra variables were more strongly related to later algebra than to later ELA; this is not entirely surprising. This indicates that prealgebra measures are better indicators of domain-specific mathematical knowledge than reading and language comprehension. There are, however, two exceptions, complex problem solving with whole numbers and geometry and measurement, that were more strongly related to later ELA than to later algebra.

One explanation is that these measures may have a meaningful reading and language comprehension component. Complex problem solving with fractions and reasoning and communicating were also good predictors of later ELA, suggesting reading and language comprehension contribute to performance on these measures.

Supplemental Regression Analyses

All the regression results presented above are based on a model that assumes the intercepts and slopes are the same for all schools and districts. To account for variability between schools and districts, Model 4 was extended for each of the outcome measures such that students are nested within schools and schools are nested within districts. The intercepts at the school and district level were modeled as random effects, as follows:

- Model 4A—Main effects only for each of the predictor variables with ELA and math scores as covariates
    - A1a, A1b, A1c, A2b, A3a, A3b, A4, AG, ELA, math
    - Random intercepts for schools and districts

Intraclass correlations (ICCs) were computed to evaluate the utility of fitting random intercepts, relative to the previously specified fixed-effects models. Table 20 shows the ICCs for each of the models. The within-school correlations and the within-district, within-school correlations are all small. These results suggest that there is insufficient variability in the effects at the school and district level to justify fitting a series of multilevel models.

Table 20. Model 4 Intraclass Correlations by Outcome Variable—Factor Scores

Intraclass correlation levels	Overall algebra	Quantitative literacy	Generalized arithmetic	Functional thinking	Geometry & statistics
Within school	0.104	0.071	0.040	0.027	0.012
Within district, within school	0.067	0.028	0.017	0.011	0.006

Although not the primary focus of the regression analyses, a subcomponent of the research question related to demographic interaction effects relates to potential differences between males and females. As such, an additional set of analyses were conducted that directly extend the specifications of these four models. The extended models, 6A, 6B, 6C, and 6D include the main effects and interaction effects for Models 5A through 5D with additional interaction effects between the demographic variables and gender. These results are presented in Appendix D.

As with Models 5A, 5B, 5C, and 5D, five outcome measures were considered: overall algebra, quantitative literacy, generalized arithmetic, functional thinking, and geometry and statistics. The differences in the coefficients for the effects modeled in Models 5A, 5B, 5C, and 5D are very small. Further, the additional effects for the interactions with gender are extremely small. This suggests that with respect to the interactions, there are no practically significant differences between males and females among the various demographic groups.

Model 5A was respecified with overall math performance as the primary predictor of interest (see Appendix D). The same five outcome measures were considered. As expected, the main effect for overall math performance is the strongest predictor, followed by the effect for ELA score. After controlling for these variables, the main effects for the demographic variables and the interaction effects are very small.

Discussion

The findings from this study have implications for the use of the SB assessment for monitoring students' readiness for high school algebra and highlight areas that might be particularly important for success in this core mathematics domain. The study also provides direction for future studies.

Implications for the SB Assessment

The main impediment to the assessment of these potentially unique relations was the small number of items administered in key content areas. To be clear, even though the total number of items for each cluster is not necessarily small, individual students received only a small number of items from each cluster. This follows from the computer-adapted nature of the test and the goal of providing an estimate of students' overall mathematical competence, not competence in specific mathematical domains.

If the SB assessment is to be used in future studies examining more nuanced relationships between earlier and later mathematics competencies (above and beyond overall mathematical competence), the number of items administered in key areas should be increased. The estimation of the effects would also be facilitated if all students received three or more items in areas (e.g., whole number or fractions arithmetic) that are thought to be critical to the preparation for later algebra.

The use of residuals and factor scores have benefits and drawbacks. The use of residuals provides the best control for overall mathematical competence, that is, assessing the relative strengths and weaknesses of individual students. However, the use of residuals likely removes too much information and is not very useful when evaluating specific strengths and later performance in algebra. This drawback is likely exacerbated by the administration of only a small number of items in each of the critical content areas.

Similarly, the estimation of subscores was hampered by too few items in some domains and was likely confounded by students' overall mathematical ability. The inclusion of Grade 5 mathematics scores in the regression models helps to control for this confound; the inclusion of ELA scores also controls for additional factors that might influence algebra performance. Without any controls for overall math, these factors could conflate general and specific competencies. That said, because the overall math score and the subscores are based on overlapping items, this can introduce bias.

At the same time, controlling for overall mathematics and ELA scores will not necessarily capture all the factors that influence performance in school (Bailey et al., 2014); hence, the effect sizes for earlier prealgebra competencies could be inflated. The use of later geometry and statistics and ELA scores as contrasts provides a means to estimate the degree of inflation—the most reliable results are those with stronger effects, relative to later algebra, although they are not optimal. A more robust measure of later geometry and statistics would have strengthened this contrast, but this would require more SB items in these areas.

Implications for Better Preparing Students for Success in Algebra

The decomposition of SB algebra items into the subdomains of quantitative literacy, generalized arithmetic, functional relations, and constraint equations makes sense in terms of students' progression through algebraic material. However, there were not enough items to construct a constraint equations measure, and the reliabilities for generalized arithmetic (reliability = .48) and functional relations (reliability = .33) were relatively low in comparison to quantitative literacy (reliability = .80) and overall algebra (reliability = .86; refer to Tables 9 and 10). The two latter measures were highly correlated (r = .96) and thus essentially are measuring the same competencies. As a result, the best outcome was overall algebra. The contrast outcome of geometry and statistics was not particularly reliable (reliability = .25), due to too few items, but was retained to provide a nonalgebra contrast outcome.

In the prediction of later overall algebra performance, the most meaningful results come from the IRT subscore analyses that controlled for earlier math and ELA scores. The comparisons of these results with the same models predicting later geometry and statistics and ELA results are important for making inferences about the specificity of the relations between early prealgebra competencies and later performance in algebra.

Interpretations of the effect sizes between earlier performance on the problem solving, complex problem solving, and reasoning and communicating measures is not as straightforward as the interpretation of performance on the number sense measures. This is because the problem-solving and reasoning measures are more complex than the concepts and procedures measures insomuch that they capture competencies that go beyond the ability to use mathematical knowledge in a problem-solving context. Performance in problem-solving contexts, as typically measured by word problems, is influenced by reading and language comprehension (Fuchs et al., 2020), students' prior knowledge as related to the context of the problem (Thevenot, 2017), and domain-general abilities, such as working memory (Geary & Widaman, 1992) and visuospatial abilities (Casey et al., 1995).

The contrast of relationships between prealgebra problem solving and reasoning competencies and later algebra and ELA scores will help to control for some of these confounds but will not likely control for all of them (Bailey et al., 2014). The current findings suggest that complex problem solving with fractions and reasoning and communicating may be more strongly related to later algebra than to later geometry and statistics and later ELA scores. This pattern also suggests that the earlier ability to use fractions and other arithmetical knowledge in the context of complex problem solving (e.g., multistep word problems) and reasoning may be important for later performance in algebra.

Limitations

The key limitation of this study is a lack of sufficient item responses, per student, within each cluster. Due to the adaptive nature of the assessment, the responses within each cluster are not missing at random. To be clear, although the entire sequence of items administered as part of an adaptive test results in missing values that can be treated as missing at random (Mislevy & Wu, 1996), it is unclear whether the item sequence for any particular item cluster satisfies this requirement. This presents enormous challenges for the creation of cluster indices, either residuals or factor scores. The residual correlations may be useful for examining unaccounted for variability and providing preliminary evidence for the presence of various clusters, but the residuals are insufficient to produce defensible indices at the student level. Stated differently, the identification of a potential factor structure versus the creation of defensible subscores for use in subsequent analyses are two distinct tasks. The data requirements are notably higher if one intends to create and provide validity evidence for any subscores. Similarly, while the IRT approach may produce scores on a metric that is more interpretable, the small number of responses for each individual leads to subscores that are not very reliable. Additionally, because the nonadministered items within each cluster are not missing at random, the resulting scores are likely biased. These issues propagate into the regression results, which are further exacerbated by the fact that individual clusters scores and residual use items in common with the overall scores and that the subscore reliabilities for the different outcome measure vary considerably. For this reason, extra care should be taken when considering inferences made on the basis of these scores.

Although there are serious limitations to this study, the value of the findings does not lie solely with the statistical inferences. Rather, the research questions and the attempted solutions with the SB data have the potential to provide guidance with respect to curricular emphasis. The results of this study do not indicate that any of the identified clusters are uniquely important with respect to predicting later algebra performance. This suggests that, based on the information that can be gleaned from the SB data, there are no obvious nor overarching shortcomings in the curricular focus in these early grades. This does not mean that the current focus cannot be improved; however, the limitations of this study should not be overlooked when considering potential curricular changes.

References

Bailey, D. H., Watts, T. W., Littlefield, A. K., & Geary, D. C. (2014). State and trait effects on individual differences in children's mathematical development. *Psychological Science, 25*(11), 2017–2026. https://doi.org/10.1177/0956797614547539

Bell, R. M., & McCaffrey, D. F. (2002). Bias reduction in standard errors for linear regression with multi-stage samples. *Survey Methodology, 28*(2), 169–182.

Birnbaum, A. (1968). Some latent trait models and their use in inferring an examinee's ability. In F. M. Lord & M. R. Novick (Eds.), *Statistical theories of mental test scores* (pp. 397–479). Addison-Wesley.

Bynner, J. M. (1997). Basic skills in adolescents' occupational preparation. *Career Development Quarterly, 45*(4), 305–321. https://doi.org/10.1002/j.2161-0045.1997.tb00536.x

Casey, M. B., Nuttall, R., Pezaris, E., & Benbow, C. P. (1995). The influence of spatial ability on gender differences in mathematics college entrance test scores across diverse samples. *Developmental Psychology*, *31*(4), 697–705. https://doi.org/10.1037/0012-1649.31.4.697

Daro, P., Hughes, G. B., & Stancavage, F. (2015). *Study of the alignment of the 2015 NAEP mathematics items at Grades 4 and 8 to the Common Core State Standards (CCSS) for Mathematics*. American Institutes for Research. https://www.air.org/sites/default/files/downloads/report/Study-of-Alignment-NAEP-Mathematics-Items-common-core-Nov-2015.pdf

Fuchs, L. S., Fuchs, D., Seethaler, P. M., & Craddock, C. (2020). Improving language comprehension to enhance word-problem solving. *Reading & Writing Quarterly*, *36*(2), 142–156. https://doi.org/10.1080/10573569.2019.1666760

Geary, D. C., Nicholas, A., Li, Y., & Sun, J. (2017). Developmental change in the influence of domain-general abilities and domain-specific knowledge on mathematics achievement: An eight-year longitudinal study. *Journal of Educational Psychology*, *109*(5), 680–693. https://doi.org/10.1037/edu0000159

Geary, D. C., & Widaman, K. F. (1992). Numerical cognition: On the convergence of componential and psychometric models. *Intelligence, 16*(1), 47–80. https://doi.org/10.1016/0160-2896(92)90025-M

Honaker, J., King, G., & Blackwell, M. (2011). Amelia II: A program for missing data. *Journal of Statistical Software, 45*(7), 1–47. https://doi.org/10.18637/jss.v045.i07

Hughes, G., Behuniak, P., Norton, S., Kitmitto, S., & Buckley, J. (2019). *NAEP validity studies NAEP panel responses to the reanalysis of TUDA mathematics scores*. American Institutes for Research. https://eric.ed.gov/?id=ED599642

Hurst, M. A., & Cordes, S. (2018). Children's understanding of fraction and decimal symbols and the notation-specific relation to pre-algebra ability. *Journal of Experimental Child Psychology*, *168*, 32–48. https://doi.org/10.1016/j.jecp.2017.12.003

Lee, K., & Bull, R. (2016). Developmental changes in working memory, updating, and math achievement. *Journal of Educational Psychology, 108*(6), 869–882. https://doi.org/10.1037/edu0000090

Mislevy, R. J., & Wu, P. K. (1996). *Missing responses and IRT ability estimation: Omits, choice, time limits, and adaptive testing* (Research Report No. RR-96-30). ETS. https://doi.org/10.1002/j.2333-8504.1996.tb01708.x

Mou, Y., Li, Y., Hoard, M. K., Nugent, L., Chu, F., Rouder, J., & Geary, D. C. (2016). Developmental foundations of children's fraction magnitude knowledge. *Cognitive Development, 39*, 141–153. https://doi.org/10.1016/j.cogdev.2016.05.002

Muraki, E. (1992). A generalized partial credit model: Application of an EM algorithm. *Applied Psychological Measurement, 16*(2), 159–176. https://doi.org/10.1177/014662169201600206

National Mathematics Advisory Panel. (2008). *Foundations for success: Final report of the National Mathematics Advisory Panel*. United States Department of Education. http://www.ed.gov/about/bdscomm/list/mathpanel/report/final-report.pdf

Schmidt, W. H., & Houang, R. T. (2012). Curricular coherence and the Common Core State Standards for Mathematics. *Educational Researcher*, *41*(8), 294–308. https://doi.org/10.3102/0013189X12464517

Siegler, R. S., & Braithwaite, D. W. (2017). Numerical development. *Annual Review of Psychology*, *68*, 187–213. https://doi.org/10.1146/annurev-psych-010416-044101

Siegler, R. S., Duncan, G. J., Davis-Kean, P. E., Duckworth, K., Claessens, A., Engel, M., Susperreguy, M. I., & Chen, M. (2012). Early predictors of high school mathematics achievement. *Psychological Science*, *23*(7), 691–697. https://doi.org/10.1177/0956797612440101

Stein, M. K., Kaufman, J. H., Sherman, M., & Hillen, A. F. (2011). Algebra: A challenge at the crossroads of policy and practice. *Review of Educational Research*, *81*(4), 453–492. https://doi.org/10.3102/0034654311423025

Thevenot, C. (2017). Arithmetic word problem solving: The role of prior knowledge. In D. C. Geary, D. B. Berch, R. Ochsendorf, & K. Man Koepke (Eds.), *Acquisition of complex arithmetic skills and higher-order mathematics concepts* (pp. 47–66). Elsevier Academic Press.

von Davier, M. (2017). *mdltm: Software for the general diagnostic model and for estimating mixtures of multidimensional discrete latent traits models* [Computer software]. ETS.

Zimba, J. (2014). The development and design of the Common Core State Standards for Mathematics. *New England Journal of Public Policy*, *26*(1), Article 10. https://scholarworks.umb.edu/cgi/viewcontent.cgi?article=1661&context=nejpp

List of Appendices

Page

Appendix A. Matching Smarter Balanced Correlates to Item Metadata

What kinds of problems do students need to be successful on in earlier grades (Table A1) to be successful in algebra (Table A2)?

To identify appropriate items, the item metadata was filtered by grade + claim+ CCSS + target model based on the metadata criteria in this section. The corresponding item categories were used to create IRT residual and IRT subscore predictor and outcome variables.

Table A1. CCSS and Metadata Used to Create Predictor Variables

Correlates of interest description	CCSS	Metadata combinations matching correlates of interest
A1a—Number sense of whole numbers	3.NBT.A.1 4.NBT.A.1 [a] 4.NBT.A.2 4.NBT.A.3	1. Claim 1 & Standard 3.NBT.A.1 2. Claim 1 & Standard 4.NBT.A.1 & TM 5a 3. Claim 1 & Standard 4.NBT.A.2 4. Claim 1 & Standard 4.NBT.A.3
A1c—Number sense of decimals and understanding of the place value system	4.NF.C.6 4.NF.C.7 5.NBT.1 [a] 5.NBT.2, 5.NBT.3a 5 NBT 3b 5.NBT.4	1. Claim 1 & Standard 4.NF.C.6 2. Claim 1 & Standard 4.NF.C.7 3. Claim 1 & Standard 5.NBT.A.2 4. Claim 1 & Standard 5.NBT.A.3a 5. Claim 1 & Standard 5.NBT.A.3b 6. Claim 1 & Standard 5.NBT.A.4
A2—Elementary modeling: one-step problems that correspond to those in tables 1 and 2, pp. 88-89, CCSSM A2—One-step word problems involving only whole numbers	3.OA.A.3 [a, b] 3.MD.A.1 [b] 3.MD.A.2 [b] 3.MD.B.3 [b] 3.MD.C.7b [b] 4.OA.A.2 [a, b] 4.MD.A.2 [a, b] 4.MD.A.3 [b] 5.MD.C.5b [b] {Any 5.NBT that match the language of A2a}	1. Claim 1 & Standard 3.OA.A.3 2. Claim 1 & Standard 3.MD.A.1 & TM 2 3. Claim 1 & Standard 3.MD.A.2 4. Claim 1 & Standard 3.MD.B.3 & TM 2 5. Claim 1 & Standard 4.OA.A.2 & TM 1a 6. Claim 1 & Standard 4.OA.A.2 & TM 1b 7. Claim 1 & Standard 4.OA.A.2 & TM 1c 8. Claim 1 & Standard 4.MD.A.3 & TM 4a 9. Claim 1 & Standard 5.MD.C.5b & TM 2b 10. Claim 1 & Standard 5.MD.C.5b & TM 2e

Correlates of interest description	CCSS	Metadata combinations matching correlates of interest
A2b—One-step word problems involving fractions	4.NF.B.3d [a, b] 4.NF.B.4c [a, b] 4.MD.A.2 [a, b] 4.MD.A.3 [b] 5.NF.B.3 [a, b] 5.NF.B.4 5.NF.B.6 [a, b] 5.NF.B.7c [a, b] 5.MD.B.2 [b] 6.NS.A.1 [a, b]	1. Claim 1 & Standard 4.NF.B.3 & TM 3a 2. Claim 1 & Standard 4.NF.B.3 & TM 3b 3. Claim 1 & Standard 4.NF.B.3 & TM 3c 4. Claim 1 & Standard 4.NF.B.3 & TM 3d 5. Claim 1 & Standard 4.NF.B.4 & TM 6a 6. Claim 1 & Standard 4.NF.B.4 & TM 6b 7. Claim 1 & Standard 5.NF.B.3 & TM 1b 8. Claim 1 & Standard 5.NF.B.3 & TM 2 9. Claim 1 & Standard 5.NF.B.4 & TM 4b 10. Claim 1 & Standard 5.NF.B.6 11. Claim 1 & Standard 5.NF.B.7 12. Claim 1 & Standard 5.MD.B.2 & TM2 13. Claim 1 & Standard 6.NS.A.1 &TM 2b 14. Claim 1 & Standard 6.NS.A.1 &TM 2c
A3—Multistep or higher-complexity word problems A3a—Multistep or higher-complexity word problems with no fractions or decimals	3.OA.D.8 [a] 3.MD.A.1 [b] 3.MD.B.3 [b] 3.MD.C.7b [b] 3.MD.C.7d [b] 3.MD.D.8 [b] 4.OA.A.2 [a, b] 4.OA.A.3 [a] 4.MD.A.2 [a, b] 4.MD.A.3 [b] 5.MD.C.5b [b] 5.MD.C.5c [b] 5.G.A.2 [b]	1. Claim 2 & any standard listed to the left (omit 4.MD.A.3 and 5.G.A.2) 2. Claim 4 & any standard listed to the left

Correlates of interest description	CCSS	Metadata combinations matching correlates of interest
A3b—Multistep or higher-complexity word problems with fractions or decimals	4.NF.B.3d [a, b] 4.NF.B.4c [a, b] 4.MD.A.2 [a, b] 4.MD.A.3 [b] 5.NF.B.3 [a, b] 5.NF.B.6 [a, b] 5.NF.B.7c [a, b] 5.MD.1 [b] 5.MD.2 [b] 5.G.A.2 [b] {any 5.NBT that match the language of A3b} 6.NS.A.1 [a, b]	1.Claim 2 & Standard 5.NF.B.3 [a, b] 5.NF.B.6 [a, b] 5.NF.B.7c [a, b] 2. Claim 2 & Standard 4.NF.B.3d [a, b] 4.NF.B.4c [a, b] 3. Claim 4 & Domain NF 4. Claim 4 & Standard 5.MD.B.2 5. Claim 4 & Standard 6.NS.A.1
A4—Reasoning and communicating Problems where students can clearly and precisely construct viable arguments to support their own reasoning and to critique the reasoning of others	3.OA.B 3.NF.A 3.NF.1 3.NF.2 3.NF.3 3.MD.A 3.MD.7 4.OA.3 4.NBT.A 4.NBT.5 4.NBT.6 4.NF.A 4.NF.1 4.NF.2 4.NF.3a 4.NF.3b 4.NF.3c 4.NF.4a 4.NF4b 4.NF.C 4.NF.7 5.NBT.2 5. NBT.7 5. NF.1 5. NF.2 5. NF.B 5. NF.3 5. NF.4 5. NF.7a 5. NF.7b 5.MD.5a 5. MD.5b	1. Claim 3 & any standard listed to the left.

Correlates of interest description	CCSS	Metadata combinations matching correlates of interest
A4—Contrast variable: Least likely to predict later outcomes in algebra	Task models in OA and NBT that are strictly procedural; geometry of shapes (nonformula based); MD items that only have students produce a graph.	1. 3.G.A.1 2. 3.G.A.2 3. 4.G.A.1 4. 4.G.A.2 5. 4.G.A.3 6. 4.MD.C.6 7. 4.MD.C.7 8. 5.G.B.3 9. 5.G.B.4 10. 5.MD.C.3 11. 5.MD.C.4

Note. HCII mapped standards codes in the data to the standards codes in column A and column B using both exact match and fuzzy match logic (to handle issues of data format—for example, codes of 3.MD.A verses 3.MD.A.1). Limitations on available item metadata resulted in corresponding limitations on the coding process. CCSS = Common Core State Standards.

[a] A central standard or CCSS organizer for the correlate in question. [b] A bucket containing tasks that match the correlate as well as tasks that don't match. For these codes, the principle for identifying tasks that match the correlate is to apply the language of the correlate as a filter. For example, the language of correlate A3a ("Multistep or higher-complexity word problems with no fractions or decimals") implies disregarding tasks in 5.G.2 if they are one-step problems or if they contain fractions or decimals. Note that the term *task* here refers not only to math problems, but also more generally to any student encounter with math that generates the desired data.

Table A2. CCSS and Metadata Used to Create Outcome Variables

Correlate of interest	CCSS	Metadata combinations matching correlates of interest
B1—Quantitative literacy (without use of variables) connecting magnitude sense to numbers and operations); identify and express relationships among quantities; represent quantities graphically. [Here we restricted the selection of quantitative literacy to that which is most applicable to the work of Algebra 1 and distant enough from the predictor variables identified in the “A” table.]	6.NS.C.5 6.NS.C.7b 6.NS.C.7c 6.NS.C.8 7.RP.A.3 [a] 7.NS.A.3 [a] 7.EE.B.3 [a] 7.G.A.1 [b] 7.G.B.6 [b] 8.EE.A.3 8.F.B.5 [a] 8.G.B.7 8.SP.A.1 8.SP.A.2 HSN-Q.A.1 HSN-Q.A.3 {higher intensity high school modeling tasks where present} {“Column 6” high school modeling tasks where present}6	1. 8.SP.A.1 2. 8.SP.A.2 3. 8.G.B.7 4. 8.F.B.5 5. 8.EE.A.3 6. 7.G.B.6 7. 7.G.A.1 8. 7.EE.B.3 9. 7.NS.A.3 10. 7.RP.A.3 11. Claim 3 & 6.NS.C.7b 12. Claim 3 & 6.NS.C.7c 13. Claim 3 & 6.NS.C.8 14. Claim 3 & 6.NS.C.6 15. Claim 3 & 6.NS.C.6b 16. Claim 3 & 6.NS.C.6c (This bucket includes both conceptual quantitative literacy and application-based problems; it doesn’t require algebraic expressions or equations.)
B2—Algebra as generalized arithmetic: writing and reading expressions and equations, transforming expressions and equations into equivalent expressions or equations using the properties of arithmetic operations and equality (linear, quadratic, exponential, or conceptual/general)	6.EE.A [a] 7.EE.A.1 [a] 8.EE.C.7b HSA-SSE.A.2 HSA-SSE.B.3a [a] HSA-SSE.B.3b [a] HSA-SSE.B.3c [a] HSA-CED.A.4	1. Claim 1 & 6.EE.A.3 2. Claim 1 & 6.EE.A.4 3. Claim 1 & 7.EE.A.1 4. Claim 1 & 8.EE.C.7b (These are all nonapplication—no context.)
B3—Algebra as functional thinking (linear, quadratic, exponential, or conceptual/general). Formulating, interpreting and using mathematical expressions, tables and graphs that refer to variable quantities and relationships between quantities	6.EE.C.9 [a] 7.RP.A.2b 7.RP.A.2c 7.EE.B.4 [b] 8.F.B.5 HSA-REI.D.11 HSF-IF [a b] HSF-BF [a b] HSF-LE.A.1 HSF-LE.A.2 HSF-LE.A.3 HSF-LE.B.5	1. 6.EE.C.9 2. 7.RP.A.2b 3. 7.RP.A.2c 4. 8.F.B.4 5. 8.F.B.5 6. 7.RP.A.2a 7. 7.RP.A.2d

Correlate of interest	CCSS	Metadata combinations matching correlates of interest
B4—Algebra as writing and solving constraint equations to solve problems in modeling scenarios (linear, quadratic, exponential, or conceptual/general)	6.EE.B.5 6.EE.B.6 [a] 6.EE.B.7 [a] 7.EE.B.4 [a] 8.EE.C.7 [b] 8.EE.C.8c [a] HSA-REI.C.6 HSA-CED.A.3 [b] HSA-REI.A.1 HSA-REI.B.3 HSA-REI.B.4 [a] HSA-REI.D.10 F-LE.A.4 {word problems matching the language of B4 not otherwise captured}	1. Claim 4 & 6.EE.B.6 3. Claim 4 & 6.EE.B.7 4. Claim 4 & 7.EE.B.4 5. Claim 4 & 7.EE.B.4a 6. Claim 4 & 8.EE.C.7 7. Claim 4 & 8.EE.C.7a 8. Claim 4 & 8.EE.C.7b 9. Claim 4 & 8.EE.C.8c (These are currently all application problems that have real world context.)
B5—Geometry & Statistics (contrast variable, least associated with Algebra 1)	1. 8.G.A.1 2. 8.G.A.2 3. 8.SP.A.4 4. 7.G.A.3 5. 7.G.A.2 6. 7.SP.A.1 7. 7.SP.A.2 8. 7.SP.B.3 9. 7.SP.B.4 10. 7.SP.C.5 11. 7.SP.C.6 12. 7.SP.C.7 13. 7.SP.C.8	1. 8.G.A.1 2. 8.G.A.2 3. 8.SP.A.4 4. 7.G.A.3 5. 7.G.A.2 6. 7.SP.A.1 7. 7.SP.A.2 8. 7.SP.B.3 9. 7.SP.B.4 10. 7.SP.C.5 11. 7.SP.C.6 12. 7.SP.C.7 13. 7.SP.C.8

Note. HCII mapped standards codes in the data to the standards codes in column A and column B using both exact match and fuzzy match logic (to handle issues of data format—for example, codes of 3.MD.A versus 3.MD.A.1). Limitations on available item metadata resulted in corresponding limitations on the coding process. CCSS = Common Core State Standards.

[a] A central standard or CCSS organizer for the correlate in question. [b] A bucket containing tasks that match the correlate as well as tasks that don't match. For these codes, the principle for identifying tasks that match the correlate is to apply the language of the correlate as a filter. For example, the language of correlate A3a ("Multistep or higher complexity word problems with no fractions or decimals") implies disregarding tasks in 5.G.2 if they are one-step problems or if they contain fractions or decimals. The term *task* here refers not only to math problems, but also more generally to any student encounter with math that generates the desired data.

Appendix B. Standardized Regressions Coefficients With Residuals

Table B1. Model 1 Standardized Regression Coefficients With Residuals by Outcome Variables

Standard	Overall algebra beta	Overall algebra SE	Quantitative literacy beta	Quantitative literacy SE	Generalized arithmetic beta	Generalized arithmetic SE	Functional thinking beta	Functional thinking SE	Geometry & statistics beta	Geometry & statistics SE
A1a—Whole numbers	-0.0005	0.0016	0.0012	0.0017	-0.0005	0.0016	**-0.0027** [a]	0.0013	-0.0009	0.0019
A1b—Fractions	0.0012	0.0010	0.0004	0.0010	0.0015	0.0011	-0.0001	0.0011	-0.0015	0.0011
A1c—Decimals	-0.0010	0.0011	-0.0003	0.0011	-0.0014	0.0010	-0.0013	0.0011	0.0008	0.0010
A2b—PS fractions	0.0002	0.0012	0.0004	0.0012	-0.0013	0.0014	0.0019	0.0013	-0.0009	0.0014
A3a—CPS whole numbers	-0.0002	0.0015	-0.0007	0.0014	-0.0004	0.0014	0.0014	0.0014	-0.0005	0.0012
A3b—CPS fractions	0.0004	0.0013	-0.0006	0.0012	0.0021	0.0012	0.0001	0.0011	-0.0001	0.0010
A4—Reasoning & communicating	0.0005	0.0011	0.0007	0.0011	0.0003	0.0011	-0.0009	0.0011	-0.0004	0.0011
AG—Geometry & measurement	0.0001	0.0014	-0.0014	0.0014	0.0016	0.0011	0.0010	0.0013	0.0004	0.0011

Note. PS = problem solving; CPS = complex problem solving.
[a] Statistically significant.

Table B2. Model 2 Standardized Regression Coefficients With Residuals by Outcome Variables

Standard	Overall algebra beta	Overall algebra SE	Quantitative literacy beta	Quantitative literacy SE	Generalized arithmetic beta	Generalized arithmetic SE	Functional thinking beta	Functional thinking SE	Geometry & statistics beta	Geometry & statistics SE
A1a—Whole numbers	-0.0005	0.0016	0.0012	0.0017	-0.0005	0.0016	**-0.0027** [a]	0.0013	-0.0009	0.0019
A1b—Fractions	0.0012	0.0010	0.0004	0.0010	0.0015	0.0011	-0.0001	0.0011	-0.0015	0.0011
A1c—Decimals	-0.0010	0.0011	-0.0003	0.0011	-0.0014	0.0010	-0.0013	0.0011	0.0008	0.0010
A2b—PS fractions	0.0002	0.0012	0.0004	0.0012	-0.0013	0.0014	0.0020	0.0013	-0.0009	0.0014
A3a—CPS whole numbers	0.0005	0.0015	-0.0006	0.0014	0.0000	0.0014	0.0024	0.0014	-0.0013	0.0012
A3b—CPS fractions	0.0004	0.0013	-0.0006	0.0012	0.0021	0.0012	0.0001	0.0011	-0.0001	0.0010
A4—Reasoning & communicating	0.0005	0.0011	0.0007	0.0011	0.0003	0.0011	-0.0009	0.0011	-0.0004	0.0011
AG—Geometry & measurement	0.0001	0.0014	-0.0014	0.0014	0.0016	0.0011	0.0010	0.0013	0.0004	0.0011
ELA	-0.0012	0.0010	-0.0002	0.0010	-0.0007	0.0010	-0.0016	0.0011	0.0014	0.0011

Note. PS = problem solving; CPS = complex problem solving; ELA = English language arts.
[a] Statistically significant.

Table B3. Model 3 Standardized Regression Coefficients With Residuals by Outcome Variables

Standard	Overall algebra beta	Overall algebra SE	Quantitative literacy beta	Quantitative literacy SE	Generalized arithmetic beta	Generalized arithmetic SE	Functional thinking beta	Functional thinking SE	Geometry & statistics beta	Geometry & statistics SE
A1a—Whole numbers	-0.0005	0.0016	0.0012	0.0017	-0.0005	0.0016	**-0.0027** [a]	0.0013	-0.0009	0.0019
A1b—Fractions	0.0012	0.0010	0.0004	0.0010	0.0015	0.0011	-0.0001	0.0011	-0.0015	0.0011
A1c—Decimals	-0.0010	0.0011	-0.0003	0.0011	-0.0014	0.0010	-0.0013	0.0011	0.0008	0.0010
A2b—PS fractions	0.0002	0.0012	0.0004	0.0012	-0.0013	0.0014	0.0019	0.0013	-0.0009	0.0014
A3a—CPS whole numbers	-0.0004	0.0015	-0.0010	0.0014	-0.0007	0.0014	0.0023	0.0014	-0.0010	0.0012
A3b—CPS fractions	0.0004	0.0013	-0.0006	0.0012	0.0021	0.0012	0.0001	0.0011	-0.0001	0.0010
A4—Reasoning & communicating	0.0005	0.0011	0.0007	0.0011	0.0003	0.0011	-0.0009	0.0011	-0.0004	0.0011
AG—Geometry & measurement	0.0001	0.0014	-0.0014	0.0014	0.0016	0.0011	0.0010	0.0013	0.0004	0.0011
Math	0.0003	0.0010	0.0005	0.0011	0.0005	0.0010	-0.0013	0.0011	0.0008	0.0011

Note. PS = problem solving; CPS = complex problem solving.
[a] Statistically significant.

Table B4. Model 4 Standardized Regression Coefficients With Residuals by Outcome Variables

Standard	Overall algebra beta	Overall algebra SE	Quantitative literacy beta	Quantitative literacy SE	Generalized arithmetic beta	Generalized arithmetic SE	Functional thinking beta	Functional thinking SE	Geometry & statistics beta	Geometry & statistics SE
A1a—Whole numbers	-0.0005	0.0016	0.0012	0.0017	-0.0005	0.0016	**-0.0027** [a]	0.0013	-0.0009	0.0019
A1b—Fractions	0.0012	0.0010	0.0004	0.0010	0.0015	0.0011	-0.0001	0.0011	-0.0015	0.0011
A1c—Decimals	-0.0010	0.0011	-0.0003	0.0011	-0.0014	0.0010	-0.0013	0.0011	0.0008	0.0010
A2b—PS fractions	0.0002	0.0012	0.0004	0.0012	-0.0013	0.0014	0.0020	0.0013	-0.0009	0.0014
A3a—CPS whole numbers	-0.0001	0.0015	-0.0009	0.0014	-0.0005	0.0014	0.0025	0.0014	-0.0012	0.0012
A3b—CPS fractions	0.0004	0.0013	-0.0006	0.0012	0.0021	0.0012	0.0001	0.0011	-0.0001	0.0010
A4—Reasoning & communicating	0.0005	0.0011	0.0007	0.0011	0.0003	0.0011	-0.0009	0.0011	-0.0004	0.0011
AG—Geometry & measurement	0.0001	0.0014	-0.0014	0.0014	0.0016	0.0011	0.0010	0.0013	0.0004	0.0011
ELA	-0.0027	0.0019	-0.0011	0.0018	-0.0019	0.0018	-0.0015	0.0019	0.0018	0.0017
Math	0.0023	0.0020	0.0013	0.0020	0.0019	0.0018	-0.0002	0.0018	-0.0005	0.0017

Note. PS = problem solving; CPS = complex problem solving.
[a] Statistically significant.

Table B5. Model 5A Standardized Regression Coefficients With Residuals by Outcome Variables: Main Effects

Standard	Overall algebra beta	Overall algebra SE	Quantitative literacy beta	Quantitative literacy SE	Generalized arithmetic beta	Generalized arithmetic SE	Functional thinking beta	Functional thinking SE	Geometry & statistics beta	Geometry & statistics SE
A4—Reasoning & communicating	**0.0050** [a]	0.0025	0.0038	0.0025	0.0017	0.0025	0.0020	0.0025	-0.0004	0.0024
Female	0.0002	0.0022	-0.0001	0.0022	-0.0008	0.0022	0.0001	0.0021	0.0015	0.0022
Hispanic or Latino	**0.0033** [a]	0.0028	**0.0028** [a]	0.0030	0.0012	0.0028	0.0019	0.0027	-0.0017	0.0028
American Indian or Alaskan Native	-0.0007	0.0157	0.0002	0.0161	-0.0004	0.0150	-0.0011	0.0156	-0.0008	0.0158
Asian	**0.0040** [a]	0.0047	**0.0040** [a]	0.0047	0.0012	0.0044	0.0014	0.0041	-0.0003	0.0041
Black or African American	0.0004	0.0055	0.0013	0.0055	0.0001	0.0053	-0.0001	0.0055	-0.0007	0.0057
Native Hawaiian or Other Pacific Islander	-0.0006	0.0154	-0.0012	0.0154	-0.0013	0.0161	0.0023	0.0171	-0.0006	0.0175
Filipino	0.0004	0.0077	0.0005	0.0076	0.0009	0.0066	-0.0001	0.0073	0.0018	0.0072
LEP	-0.0018	0.0030	-0.0022	0.0031	-0.0009	0.0029	-0.0006	0.0030	-0.0016	0.0029
504 plan	-0.0005	0.0099	-0.0008	0.0098	0.0006	0.0096	0.0002	0.0112	0.0004	0.0101
Economic disadvantage	-0.0021	0.0026	-0.0010	0.0026	-0.0008	0.0028	-0.0013	0.0026	-0.0002	0.0024
ELA	-0.0023	0.0012	-0.0017	0.0011	-0.0011	0.0013	-0.0006	0.0013	-0.0007	0.0013

Note. LEP = limited English proficiency; ELA = English language arts.
[a] Statistically significant.

Table B6. Model 5A Standardized Regression Coefficients With Residuals by Outcome Variables: Interaction Effects

Standard	Overall algebra beta	Overall algebra SE	Quantitative literacy beta	Quantitative literacy SE	Generalized arithmetic beta	Generalized arithmetic SE	Functional thinking beta	Functional thinking SE	Geometry & statistics beta	Geometry & statistics SE
A4: female	-0.0008	0.0022	-0.0017	0.0021	0.0016	0.0024	-0.0012	0.0021	0.0007	0.0021
A4: Hispanic or Latino	-0.0020	0.0030	-0.0017	0.0031	-0.0019	0.0029	-0.0003	0.0032	0.0033	0.0027
A4: American Indian or Alaskan Native	-0.0017	0.0159	-0.0020	0.0158	0.0009	0.0152	-0.0014	0.0140	0.0016	0.0164
A4: Asian	-0.0002	0.0045	-0.0003	0.0044	0.0005	0.0040	0.0001	0.0040	0.0003	0.0045
A4: Black or African American	-0.0005	0.0049	0.0002	0.0050	-0.0014	0.0049	0.0002	0.0056	0.0011	0.0049
A4: Native Hawaiian or Other Pacific Islander	0.0004	0.0168	0.0005	0.0169	-0.0015	0.0135	0.0011	0.0160	0.0017	0.0142
A4: Filipino	-0.0020	0.0076	-0.0002	0.0076	**-0.0028** [a]	0.0070	-0.0015	0.0073	0.0005	0.0063
A4: LEP	0.0000	0.0030	-0.0002	0.0029	-0.0005	0.0027	0.0000	0.0029	-0.0016	0.0028
A4: 504 plan	-0.0018	0.0100	**-0.0022** [a]	0.0099	-0.0007	0.0105	0.0001	0.0107	0.0006	0.0100
A4: Economic disadvantage	-0.0023	0.0028	-0.0001	0.0028	-0.0003	0.0026	-0.0026	0.0027	**-0.0040** [a]	0.0025

Note. LEP = limited English proficiency.
[a] Statistically significant.

Table B7. Model 5B Standardized Regression Coefficients With Residuals by Outcome Variables: Main Effects

Standard	Overall algebra beta	Overall algebra SE	Quantitative literacy beta	Quantitative literacy SE	Generalized arithmetic beta	Generalized arithmetic SE	Functional thinking beta	Functional thinking SE	Geometry & statistics beta	Geometry & statistics SE
A3b—CPS fractions	0.0000	0.0026	0.0019	0.0025	0.0001	0.0027	-0.0036	0.0026	0.0011	0.0026
Female	0.0001	0.0022	0.0000	0.0022	-0.0011	0.0022	0.0003	0.0021	0.0013	0.0022
Hispanic or Latino	**0.0036** [a]	0.0028	**0.0031** [a]	0.0030	0.0014	0.0028	0.0021	0.0027	-0.0023	0.0028
American Indian or Alaskan Native	-0.0004	0.0156	0.0006	0.0161	-0.0005	0.0150	-0.0009	0.0156	-0.0010	0.0158
Asian	**0.0042** [a]	0.0047	**0.0040** [a]	0.0047	0.0013	0.0044	0.0016	0.0042	-0.0006	0.0041
Black or African American	0.0005	0.0055	0.0011	0.0055	0.0004	0.0053	0.0001	0.0055	-0.0009	0.0057
Native Hawaiian or Other Pacific Islander	-0.0007	0.0154	-0.0013	0.0155	-0.0009	0.0161	0.0022	0.0169	-0.0009	0.0175
Filipino	0.0007	0.0078	0.0004	0.0076	0.0014	0.0066	0.0003	0.0074	0.0018	0.0072
LEP	-0.0019	0.0030	-0.0023	0.0031	-0.0007	0.0029	-0.0006	0.0030	-0.0014	0.0029
504 plan	-0.0001	0.0100	-0.0003	0.0098	0.0006	0.0096	0.0002	0.0112	0.0004	0.0101
Economic disadvantage	-0.0017	0.0026	-0.0009	0.0026	-0.0005	0.0028	-0.0009	0.0026	0.0003	0.0024
ELA	-0.0023	0.0012	-0.0017	0.0011	-0.0011	0.0013	-0.0006	0.0013	-0.0007	0.0013

Note. CPS = complex problem solving; LEP = limited English proficiency; ELA = English language arts.
[a] Statistically significant.

Table B8. Model 5B Standardized Regression Coefficients With Residuals by Outcome Variables: Interaction Effects

Standard	Overall algebra beta	Overall algebra SE	Quantitative literacy beta	Quantitative literacy SE	Generalized arithmetic beta	Generalized arithmetic SE	Functional thinking beta	Functional thinking SE	Geometry & statistics beta	Geometry & statistics SE
A3b: female	-0.0025	0.0021	-0.0023	0.0021	-0.0014	0.0021	-0.0001	0.0022	-0.0004	0.0021
A3b: Hispanic or Latino	0.0005	0.0028	0.0002	0.0029	-0.0019	0.0029	0.0026	0.0029	-0.0012	0.0031
A3b: American Indian or Alaskan Native	-0.0003	0.0166	-0.0001	0.0174	0.0002	0.0141	-0.0004	0.0158	0.0005	0.0160
A3b: Asian	0.0014	0.0042	-0.0006	0.0042	0.0018	0.0040	0.0023	0.0043	-0.0016	0.0043
A3b: Black or African American	-0.0003	0.0052	-0.0015	0.0054	-0.0001	0.0050	**0.0022** [a]	0.0049	0.0007	0.0061
A3b: Native Hawaiian or Other Pacific Islander	0.0003	0.0166	0.0002	0.0187	0.0007	0.0147	0.0011	0.0160	0.0007	0.0153
A3b: Filipino	-0.0005	0.0071	-0.0012	0.0075	-0.0007	0.0080	0.0013	0.0072	0.0013	0.0079
A3b: LEP	-0.0012	0.0027	-0.0021	0.0026	0.0012	0.0027	-0.0006	0.0031	-0.0008	0.0028
A3b: 504 plan	0.0006	0.0097	0.0012	0.0100	-0.0013	0.0099	0.0002	0.0108	0.0007	0.0096
A3b: Economic disadvantage	0.0026	0.0026	0.0008	0.0024	**0.0045** [a]	0.0025	0.0008	0.0025	0.0002	0.0026

Note. LEP = limited English proficiency.
[a] Statistically significant.

Table B9. Model 5C Standardized Regression Coefficients With Residuals by Outcome Variables: Main Effects

Standard	Overall algebra beta	Overall algebra SE	Quantitative literacy beta	Quantitative literacy SE	Generalized arithmetic beta	Generalized arithmetic SE	Functional thinking beta	Functional thinking SE	Geometry & statistics beta	Geometry & statistics SE
A1b—Fractions	0.0021	0.0024	0.0011	0.0025	0.0027	0.0024	-0.0015	0.0025	-0.0038	0.0025
Female	0.0003	0.0022	0.0002	0.0022	-0.0012	0.0022	0.0004	0.0021	0.0014	0.0022
Hispanic or Latino	**0.0035** [a]	0.0028	**0.0030** [a]	0.0030	0.0016	0.0028	0.0019	0.0027	-0.0020	0.0028
American Indian or Alaskan Native	-0.0004	0.0156	0.0006	0.0160	-0.0005	0.0150	-0.0009	0.0156	-0.0011	0.0158
Asian	**0.0041** [a]	0.0047	**0.0042** [a]	0.0047	0.0011	0.0044	0.0013	0.0041	-0.0005	0.0041
Black or African American	0.0008	0.0055	0.0013	0.0055	0.0006	0.0053	-0.0001	0.0055	-0.0009	0.0057
Native Hawaiian or Other Pacific Islander	-0.0008	0.0154	-0.0013	0.0155	-0.0012	0.0161	0.0021	0.0170	-0.0010	0.0174
Filipino	0.0007	0.0078	0.0005	0.0076	0.0014	0.0066	0.0002	0.0073	0.0016	0.0072
LEP	-0.0018	0.0030	-0.0021	0.0031	-0.0009	0.0029	-0.0005	0.0030	-0.0013	0.0029
504 plan	-0.0001	0.0100	-0.0005	0.0098	0.0009	0.0096	0.0002	0.0112	0.0003	0.0101
Economic disadvantage	-0.0019	0.0026	-0.0010	0.0026	-0.0009	0.0028	-0.0010	0.0026	0.0004	0.0024
ELA	-0.0023	0.0012	-0.0017	0.0011	-0.0011	0.0013	-0.0006	0.0013	-0.0007	0.0013

Note. LEP = limited English proficiency; ELA = English language arts.
[a] Statistically significant.

Table B10. Model 5C Standardized Regression Coefficients With Residuals by Outcome Variables: Interaction Effects

Standard	Overall algebra beta	Overall algebra SE	Quantitative literacy beta	Quantitative literacy SE	Generalized arithmetic beta	Generalized arithmetic SE	Functional thinking beta	Functional thinking SE	Geometry & statistics beta	Geometry & statistics SE
A1b: female	-0.0001	0.0021	0.0001	0.0022	-0.0022	0.0020	0.0022	0.0021	0.0003	0.0021
A1b: Hispanic or Latino	-0.0011	0.0028	-0.0013	0.0029	0.0012	0.0029	-0.0011	0.0028	0.0022	0.0030
A1b: American Indian or Alaskan Native	-0.0006	0.0168	0.0000	0.0173	0.0000	0.0142	0.0000	0.0153	-0.0001	0.0166
A1b: Asian	0.0006	0.0034	0.0012	0.0035	-0.0003	0.0035	-0.0006	0.0042	-0.0009	0.0040
A1b: Black or African American	0.0018	0.0051	0.0010	0.0053	0.0018	0.0054	0.0010	0.0048	0.0003	0.0051
A1b: Native Hawaiian or Other Pacific Islander	-0.0010	0.0143	-0.0001	0.0159	-0.0017	0.0171	-0.0004	0.0154	-0.0002	0.0162
A1b: Filipino	-0.0004	0.0071	-0.0007	0.0072	-0.0001	0.0069	0.0004	0.0077	-0.0005	0.0066
A1b: LEP	0.0003	0.0027	0.0007	0.0029	-0.0011	0.0029	0.0005	0.0027	0.0001	0.0026
A1b: 501b plan	0.0003	0.0117	-0.0008	0.0123	0.0015	0.0106	0.0001	0.0110	0.0001	0.0103
A1b: Economic disadvantage	-0.0007	0.0026	-0.0004	0.0027	-0.0006	0.0026	0.0001	0.0023	0.0009	0.0026

Note. LEP = limited English proficiency.
[a] Statistically significant.

Table B11. Model 5D Standardized Regression Coefficients With Residuals by Outcome Variables: Main Effects

Standard	Overall algebra beta	Overall algebra SE	Quantitative literacy beta	Quantitative literacy SE	Generalized arithmetic beta	Generalized arithmetic SE	Functional thinking beta	Functional thinking SE	Geometry & statistics beta	Geometry & statistics SE
A3a—CPS whole numbers	-0.0008	0.0027	-0.0002	0.0028	-0.0032	0.0027	-0.0006	0.0025	-0.0048	0.0027
Female	0.0004	0.0022	0.0003	0.0022	-0.0010	0.0022	0.0002	0.0021	0.0015	0.0022
Hispanic or Latino	**0.0038** [a]	0.0028	**0.0033** [a]	0.0030	0.0017	0.0028	0.0019	0.0027	-0.0016	0.0028
American Indian or Alaskan Native	-0.0006	0.0156	0.0003	0.0161	-0.0005	0.0150	-0.0008	0.0156	-0.0009	0.0158
Asian	**0.0039** [a]	0.0047	**0.0039** [a]	0.0047	0.0012	0.0044	0.0013	0.0041	-0.0001	0.0041
Black or African American	0.0010	0.0055	0.0015	0.0055	0.0007	0.0054	-0.0001	0.0055	-0.0003	0.0057
Native Hawaiian or Other Pacific Islander	-0.0007	0.0154	-0.0013	0.0155	-0.0011	0.0162	0.0021	0.0170	-0.0010	0.0174
Filipino	0.0008	0.0078	0.0007	0.0076	0.0017	0.0066	-0.0001	0.0074	0.0021	0.0072
LEP	-0.0015	0.0030	-0.0020	0.0031	-0.0005	0.0029	-0.0003	0.0030	-0.0014	0.0029
504 plan	-0.0002	0.0100	-0.0002	0.0098	0.0005	0.0096	0.0000	0.0112	0.0004	0.0101
Economic disadvantage	-0.0019	0.0027	-0.0012	0.0026	-0.0007	0.0028	-0.0010	0.0026	0.0002	0.0024
ELA	-0.0023	0.0012	-0.0017	0.0011	-0.0011	0.0013	-0.0006	0.0013	-0.0007	0.0013

Note. CPS = complex problem solving; LEP = limited English proficiency; ELA = English language arts.
[a] Statistically significant.

Table B12. Model 5D Standardized Regression Coefficients With Residuals by Outcome Variables: Interaction Effects

Standard	Overall algebra beta	Overall algebra SE	Quantitative literacy beta	Quantitative literacy SE	Generalized arithmetic beta	Generalized arithmetic SE	Functional thinking beta	Functional thinking SE	Geometry & statistics beta	Geometry & statistics SE
A3a: female	0.0006	0.0023	0.0014	0.0024	0.0000	0.0022	-0.0004	0.0023	0.0013	0.0022
A3a: Hispanic or Latino	0.0014	0.0032	0.0021	0.0030	0.0013	0.0031	-0.0002	0.0041	**0.0050** [a]	0.0029
A3a: American Indian or Alaskan Native	-0.0010	0.0192	-0.0018	0.0192	0.0002	0.0176	0.0001	0.0162	0.0009	0.0163
A3a: Asian	-0.0007	0.0043	-0.0010	0.0048	0.0005	0.0051	-0.0003	0.0050	0.0019	0.0043
A3a: Black or African American	0.0023	0.0059	0.0018	0.0057	0.0016	0.0057	0.0002	0.0061	**0.0037** [a]	0.0059
A3a: Native Hawaiian or Other Pacific Islander	0.0001	0.0190	-0.0001	0.0174	-0.0004	0.0164	-0.0001	0.0188	-0.0003	0.0160
A3a: Filipino	0.0003	0.0073	0.0008	0.0069	0.0014	0.0085	-0.0013	0.0082	0.0021	0.0087
A3a: LEP	0.0022	0.0029	0.0014	0.0028	0.0018	0.0030	0.0017	0.0036	-0.0007	0.0030
A3a: 504 plan	0.0001	0.0110	0.0010	0.0117	-0.0013	0.0124	-0.0008	0.0126	0.0005	0.0117
A3a: Economic disadvantage	-0.0011	0.0024	-0.0018	0.0025	0.0008	0.0026	0.0002	0.0029	-0.0011	0.0027

Note. LEP = limited English proficiency.
[a] Statistically significant.

Table B13. *R*-squared Statistics for Regression Models With Residuals by Outcome Variables

Model	Overall algebra	Quantitative literacy	Generalized arithmetic	Functional thinking	Geometry & statistics
1	0.000006	0.000005	0.000015	0.000020	0.000008
2	0.000007	0.000005	0.000015	0.000020	0.000008
3	0.000006	0.000005	0.000015	0.000020	0.000008
4	0.000009	0.000005	0.000016	0.000020	0.000009
5A	0.000032	0.000030	0.000021	0.000018	0.000025
5B	0.000027	0.000029	0.000026	0.000018	0.000020
5C	0.000027	0.000024	0.000020	0.000014	0.000021
5D	0.000028	0.000034	0.000016	0.000017	0.000030

Appendix C. Standardized Regression Coefficients With Factor Scores

Table C1. Model 1 Standardized Regression Coefficients With Factor Scores by Outcome Variables

Standard	Overall algebra beta [a]	Overall algebra SE	Quantitative literacy beta [a]	Quantitative literacy SE	Generalized arithmetic beta [a]	Generalized arithmetic SE	Functional thinking beta [a]	Functional thinking SE	Geometry & statistics beta [a]	Geometry & statistics SE
A1a—Whole numbers	**0.059**	0.001	**0.056**	0.001	**0.037**	0.002	**0.023**	0.002	**0.025**	0.002
A1b—Fractions	**0.159**	0.001	**0.158**	0.001	**0.058**	0.002	**0.074**	0.002	**0.084**	0.002
A1c—Decimals	**0.057**	0.001	**0.053**	0.001	**0.038**	0.001	**0.027**	0.001	**0.021**	0.001
A2b—PS fractions	**0.071**	0.001	**0.069**	0.001	**0.033**	0.002	**0.028**	0.001	**0.042**	0.002
A3a—CPS whole numbers	**0.095**	0.001	**0.087**	0.001	**0.049**	0.002	**0.046**	0.002	**0.049**	0.002
A3b—CPS fractions	**0.198**	0.001	**0.192**	0.001	**0.093**	0.002	**0.103**	0.002	**0.103**	0.002
A4—Reasoning & communicating	**0.386**	0.002	**0.367**	0.002	**0.208**	0.004	**0.207**	0.003	**0.188**	0.003
AG—Geometry & measurement	**0.080**	0.001	**0.069**	0.001	**0.068**	0.002	**0.051**	0.001	**0.042**	0.001

Note. PS = problem solving; CPS = complex problem solving.
[a] All values in the beta columns are statistically significant.

Table C2. Model 2 Standardized Regression Coefficients With Factor Scores by Outcome Variables

Standard	Overall algebra beta [a]	Overall algebra SE	Quantitative literacy beta [a]	Quantitative literacy SE	Generalized arithmetic beta [a]	Generalized arithmetic SE	Functional thinking beta [a]	Functional thinking SE	Geometry & statistics beta [a]	Geometry & statistics SE
A1a—Whole numbers	**0.050**	0.001	**0.048**	0.001	**0.030**	0.002	**0.016**	0.002	**0.018**	0.002
A1b—Fractions	**0.127**	0.001	**0.131**	0.001	**0.035**	0.002	**0.048**	0.002	**0.061**	0.002
A1c—Decimals	**0.050**	0.001	**0.046**	0.001	**0.032**	0.001	**0.021**	0.001	**0.015**	0.001
A2b—PS fractions	**0.059**	0.001	**0.060**	0.001	**0.024**	0.002	**0.019**	0.001	**0.034**	0.002
A3a—CPS whole numbers	**0.062**	0.001	**0.059**	0.001	**0.026**	0.001	**0.020**	0.002	**0.025**	0.002
A3b—CPS fractions	**0.146**	0.001	**0.149**	0.001	**0.056**	0.002	**0.061**	0.002	**0.066**	0.002
A4—Reasoning & communicating	**0.283**	0.002	**0.281**	0.002	**0.134**	0.003	**0.124**	0.003	**0.114**	0.003
AG—Geometry & measurement	**0.056**	0.001	**0.049**	0.001	**0.050**	0.001	**0.032**	0.001	**0.025**	0.001
ELA	**0.252**	0.002	**0.213**	0.002	**0.181**	0.003	**0.204**	0.002	**0.182**	0.002

Note. PS = problem solving; CPS = complex problem solving; ELA = English language arts.
[a] All values in the beta columns are statistically significant.

Table C3. Model 3 Standardized Regression Coefficients With Factor Scores by Outcome Variables

Standard	Overall algebra beta	Overall algebra SE	Quantitative literacy beta	Quantitative literacy SE	Generalized arithmetic beta	Generalized arithmetic SE	Functional thinking beta	Functional thinking SE	Geometry & statistics beta	Geometry & statistics SE
A1a—Whole numbers	**0.040** [a]	0.001	**0.038** [a]	0.001	**0.025** [a]	0.002	**0.012** [a]	0.002	**0.016** [a]	0.002
A1b—Fractions	**0.020** [a]	0.001	**0.030** [a]	0.001	**-0.029** [a]	0.002	**-0.006** [a]	0.002	**0.021** [a]	0.002
A1c—Decimals	**0.010** [a]	0.001	**0.009** [a]	0.001	**0.008** [a]	0.001	-0.001	0.001	-0.001	0.001
A2b—PS fractions	**0.015** [a]	0.001	**0.018** [a]	0.001	-0.003	0.002	**-0.004** [a]	0.001	**0.017** [a]	0.002
A3a—CPS whole numbers	**0.049** [a]	0.001	**0.045** [a]	0.002	**0.020** [a]	0.002	**0.020** [a]	0.002	**0.028** [a]	0.002
A3b—CPS fractions	**0.064** [a]	0.002	**0.069** [a]	0.002	**0.009** [a]	0.002	**0.026** [a]	0.002	**0.043** [a]	0.002
A4—Reasoning & communicating	**0.160** [a]	0.002	**0.159** [a]	0.002	**0.066** [a]	0.004	**0.077** [a]	0.003	**0.086** [a]	0.002
AG—Geometry & measurement	**0.016** [a]	0.001	**0.010** [a]	0.001	**0.027** [a]	0.001	**0.014** [a]	0.001	**0.013** [a]	0.001
Math	**0.544** [a]	0.004	**0.502** [a]	0.005	**0.340** [a]	0.004	**0.313** [a]	0.004	**0.245** [a]	0.004

Note. PS = problem solving; CPS = complex problem solving.
[a] Statistically significant.

Table C4. Model 4 Standardized Regression Coefficients With Factor Scores by Outcome Variables

Standard	Overall algebra beta	Overall algebra SE	Quantitative literacy beta	Quantitative literacy SE	Generalized arithmetic beta	Generalized arithmetic SE	Functional thinking beta	Functional thinking SE	Geometry & statistics beta	Geometry & statistics SE
A1a—Whole numbers	**0.037** [a]	0.001	**0.036** [a]	0.001	**0.022** [a]	0.002	**0.009** [a]	0.002	**0.013** [a]	0.002
A1b—Fractions	**0.023** [a]	0.001	**0.032** [a]	0.001	**-0.027** [a]	0.002	**-0.004** [a]	0.002	**0.023** [a]	0.001
A1c—Decimals	**0.013** [a]	0.001	**0.011** [a]	0.001	**0.010** [a]	0.001	**0.002** [a]	0.001	0.002	0.001
A2b—PS fractions	**0.017** [a]	0.001	**0.019** [a]	0.001	-0.001	0.002	-0.002	0.001	**0.019** [a]	0.002
A3a—CPS whole numbers	**0.035** [a]	0.001	**0.034** [a]	0.001	**0.009** [a]	0.002	**0.006** [a]	0.002	**0.015** [a]	0.002
A3b—CPS fractions	**0.053** [a]	0.001	**0.060** [a]	0.002	0.000	0.002	**0.015** [a]	0.002	**0.033** [a]	0.002
A4—Reasoning & communicating	**0.130** [a]	0.002	**0.135** [a]	0.002	**0.043** [a]	0.003	**0.048** [a]	0.003	**0.059** [a]	0.002
AG—Geometry & measurement	**0.010** [a]	0.001	**0.006** [a]	0.001	**0.023** [a]	0.001	**0.009** [a]	0.001	**0.008** [a]	0.001
ELA	**0.165** [a]	0.002	**0.130** [a]	0.002	**0.129** [a]	0.003	**0.160** [a]	0.002	**0.151** [a]	0.002
Math	**0.454** [a]	0.005	**0.430** [a]	0.005	**0.269** [a]	0.004	**0.225** [a]	0.003	**0.163** [a]	0.004

Note. PS = problem solving; CPS = complex problem solving; ELA = English language arts.
[a] Statistically significant.

Table C5. Model 5A Standardized Regression Coefficients With Factor Scores by Outcome Variables: Main Effects

Standard	Overall algebra beta	Overall algebra SE	Quantitative literacy beta	Quantitative literacy SE	Generalized arithmetic beta	Generalized arithmetic SE	Functional thinking beta	Functional thinking SE	Geometry & statistics beta	Geometry & statistics SE
A4—Reasoning & communicating	**0.435** [a]	0.003	**0.442** [a]	0.003	**0.142** [a]	0.006	**0.175** [a]	0.003	**0.224** [a]	0.004
Female	**-0.050** [a]	0.001	**-0.054** [a]	0.002	**-0.004** [a]	0.002	**0.007** [a]	0.002	**-0.043** [a]	0.002
Hispanic or Latino	**-0.031** [a]	0.005	**-0.040** [a]	0.005	**0.036** [a]	0.007	0.007	0.005	**-0.064** [a]	0.005
American Indian or Alaskan Native	**-0.008** [a]	0.012	**-0.007** [a]	0.012	**-0.002** [a]	0.018	**-0.006** [a]	0.016	**-0.005** [a]	0.016
Asian	**0.090** [a]	0.014	**0.084** [a]	0.014	**0.065** [a]	0.011	**0.047** [a]	0.010	**0.031** [a]	0.007
Black or African American	**-0.040** [a]	0.008	**-0.046** [a]	0.007	**0.004** [a]	0.011	**-0.006** [a]	0.009	**-0.034** [a]	0.008
Native Hawaiian or Other Pacific Islander	-0.001	0.013	-0.002	0.013	0.004	0.018	0.001	0.017	**-0.004** [a]	0.015
Filipino	**0.017** [a]	0.009	**0.016** [a]	0.009	**0.014** [a]	0.012	**0.018** [a]	0.011	**-0.009** [a]	0.008
LEP	**-0.038** [a]	0.008	**-0.051** [a]	0.008	0.033	0.008	**-0.020** [a]	0.007	**-0.042** [a]	0.004
504 plan	**-0.006** [a]	0.008	**-0.004** [a]	0.008	**-0.006** [a]	0.010	**-0.004** [a]	0.010	-0.001	0.011
Economic disadvantage	**-0.051** [a]	0.006	**-0.059** [a]	0.006	**0.013** [a]	0.007	**-0.001** [a]	0.005	**-0.063** [a]	0.005
ELA	**0.363** [a]	0.002	**0.325** [a]	0.002	**0.220** [a]	0.004	**0.242** [a]	0.003	**0.223** [a]	0.002

Note. LEP = limited English proficiency; ELA = English language arts.
[a] Statistically significant.

Table C6. Model 5A Standardized Regression Coefficients With Factor Scores by Outcome Variables: Interaction Effects

Standard	Overall algebra beta	Overall algebra SE	Quantitative literacy beta	Quantitative literacy SE	Generalized arithmetic beta	Generalized arithmetic SE	Functional thinking beta	Functional thinking SE	Geometry & statistics beta	Geometry & statistics SE
A4: female	**-0.034** [a]	0.001	**-0.026** [a]	0.002	**-0.029** [a]	0.002	**-0.017** [a]	0.002	**-0.033** [a]	0.002
A4: Hispanic or Latino	**-0.007** [a]	0.003	**-0.016** [a]	0.003	**0.042** [a]	0.004	**0.006** [a]	0.004	**-0.040** [a]	0.003
A4: American Indian or Alaskan Native	-0.002	0.010	**-0.003** [a]	0.011	0.003	0.015	-0.001	0.017	-0.004	0.017
A4: Asian	0.001	0.007	0.001	0.007	-0.001	0.008	-0.004	0.007	**0.015** [a]	0.007
A4: Black or African American	**-0.011** [a]	0.006	**-0.023** [a]	0.006	**0.029** [a]	0.007	**0.009** [a]	0.006	**-0.015** [a]	0.006
A4: Native Hawaiian or Other Pacific Islander	-0.001	0.010	**-0.003** [a]	0.010	0.003	0.016	0.001	0.016	0.000	0.018
A4: Filipino	**-0.003** [a]	0.006	**-0.003** [a]	0.006	**-0.005** [a]	0.008	**0.005** [a]	0.008	-0.001	0.008
A4: LEP	**-0.028** [a]	0.005	**-0.044** [a]	0.006	**0.064** [a]	0.004	**-0.018** [a]	0.004	**-0.028** [a]	0.003
A4: 504 plan	-0.001	0.007	-0.001	0.007	0.000	0.012	-0.002	0.011	-0.002	0.011
A4: Economic disadvantage	**-0.007** [a]	0.003	**-0.020** [a]	0.003	**0.046** [a]	0.004	**0.020** [a]	0.004	**-0.047** [a]	0.003

Note. LEP = limited English proficiency.
[a] Statistically significant.

Table C7. Model 5B Standardized Regression Coefficients With Factor Scores by Outcome Variables: Main Effects

Standard	Overall algebra beta	Overall algebra SE	Quantitative literacy beta	Quantitative literacy SE	Generalized arithmetic beta	Generalized arithmetic SE	Functional thinking beta	Functional thinking SE	Geometry & statistics beta	Geometry & statistics SE
A3b—CPS fractions	**0.260** [a]	0.003	**0.268** [a]	0.003	**0.069** [a]	0.003	**0.100** [a]	0.002	**0.145** [a]	0.001
Female	**-0.074** [a]	0.002	**-0.076** [a]	0.002	**-0.020** [a]	0.002	**-0.003** [a]	0.001	**-0.054** [a]	0.001
Hispanic or Latino	**-0.041** [a]	0.003	**-0.051** [a]	0.003	**0.036** [a]	0.003	0.004	0.002	**-0.071** [a]	0.001
American Indian or Alaskan Native	**-0.009** [a]	0.018	**-0.008** [a]	0.019	**0.000** [a]	0.017	**-0.005** [a]	0.011	**-0.007** [a]	0.007
Asian	**0.106** [a]	0.003	**0.099** [a]	0.003	**0.069** [a]	0.003	**0.052** [a]	0.002	**0.043** [a]	0.001
Black or African American	**-0.050** [a]	0.007	**-0.057** [a]	0.007	**0.005** [a]	0.006	**-0.008** [a]	0.004	**-0.042** [a]	0.003
Native Hawaiian or Other Pacific Islander	-0.003	0.018	-0.004	0.018	0.003	0.016	-0.001	0.011	**-0.007** [a]	0.007
Filipino	**0.018** [a]	0.006	**0.016** [a]	0.007	**0.013** [a]	0.006	**0.019** [a]	0.004	**-0.008** [a]	0.002
LEP	**-0.038** [a]	0.004	**-0.050** [a]	0.004	**0.029** [a]	0.004	**-0.018** [a]	0.003	**-0.042** [a]	0.002
504 plan	**-0.009** [a]	0.011	**-0.006** [a]	0.011	**-0.008** [a]	0.010	**-0.006** [a]	0.007	-0.001	0.004
Economic disadvantage	**-0.062** [a]	0.003	**-0.070** [a]	0.003	**0.008** [a]	0.002	**-0.004** [a]	0.002	**-0.069** [a]	0.001
ELA	**0.508** [a]	0.001	**0.467** [a]	0.001	**0.293** [a]	0.001	**0.305** [a]	0.000	**0.278** [a]	0.000

Note. CPS = complex problem solving; LEP = limited English proficiency; ELA = English language arts.
[a] Statistically significant.

Table C8. Model 5B Standardized Regression Coefficients With Factor Scores by Outcome Variables: Interaction Effects

Standard	Overall algebra beta	Overall algebra SE	Quantitative literacy beta	Quantitative literacy SE	Generalized arithmetic beta	Generalized arithmetic SE	Functional thinking beta	Functional thinking SE	Geometry & statistics beta	Geometry & statistics SE
A3b: Female	**-0.037** [a]	0.002	**-0.029** [a]	0.002	**-0.035** [a]	0.002	**-0.017** [a]	0.001	**-0.033** [a]	0.001
A3b: Hispanic or Latino	**-0.009** [a]	0.003	**-0.017** [a]	0.003	**0.039** [a]	0.003	0.005	0.002	**-0.038** [a]	0.001
A3b: American Indian or Alaskan Native	-0.001	0.018	-0.002	0.019	**0.006** [a]	0.016	0.001	0.011	-0.004	0.007
A3b: Asian	0.002	0.004	0.002	0.004	-0.003	0.004	-0.003	0.003	**0.016** [a]	0.002
A3b: Black or African American	**-0.013** [a]	0.007	**-0.026** [a]	0.007	**0.031** [a]	0.006	**0.010** [a]	0.004	**-0.020** [a]	0.003
A3b: Native Hawaiian or Other Pacific Islander	-0.002	0.018	-0.003	0.019	0.002	0.017	-0.001	0.011	-0.003	0.007
A3b: Filipino	**-0.004** [a]	0.008	**-0.004** [a]	0.008	**-0.007** [a]	0.007	0.003	0.005	-0.001	0.003
A3b: LEP	**-0.022** [a]	0.004	**-0.037** [a]	0.004	**0.062** [a]	0.003	**-0.012** [a]	0.002	**-0.028** [a]	0.001
A3b: 504 plan	-0.002	0.012	-0.001	0.012	-0.002	0.011	-0.003	0.007	-0.002	0.004
A3b: Economic disadvantage	-0.005	0.003	**-0.015**	0.003	**0.036** [a]	0.003	**0.018** [a]	0.002	**-0.037** [a]	0.001

Note. LEP = limited English proficiency.
[a] Statistically significant.

Table C9. Model 5C Standardized Regression Coefficients With Factor Scores by Outcome Variables: Main Effects

Standard	Overall algebra beta	Overall algebra SE	Quantitative literacy beta	Quantitative literacy SE	Generalized arithmetic beta	Generalized arithmetic SE	Functional thinking beta	Functional thinking SE	Geometry & statistics beta	Geometry & statistics SE
A1b—Fractions	**0.218**	0.003	**0.228**	0.003	**0.046**	0.004	**0.077**	0.003	**0.120**	0.003
Female	**-0.052**	0.002	**-0.058**	0.002	**-0.005**	0.002	**0.008**	0.002	**-0.034**	0.002
Hispanic or Latino	**-0.039** [a]	0.006	**-0.046** [a]	0.005	**0.016** [a]	0.007	0.001	0.005	**-0.052** [a]	0.006
American Indian or Alaskan Native	**-0.008** [a]	0.013	**-0.007** [a]	0.012	**-0.002** [a]	0.018	**-0.005** [a]	0.016	**-0.003** [a]	0.015
Asian	**0.112** [a]	0.013	**0.106** [a]	0.014	**0.073** [a]	0.012	**0.058** [a]	0.009	**0.037** [a]	0.010
Black or African American	**-0.044** [a]	0.009	**-0.044** [a]	0.008	**-0.012** [a]	0.010	**-0.015** [a]	0.009	**-0.032** [a]	0.008
Native Hawaiian or Other Pacific Islander	-0.003	0.014	-0.003	0.014	0.001	0.018	0.000	0.017	**-0.006** [a]	0.015
Filipino	**0.022** [a]	0.011	**0.020** [a]	0.010	**0.017** [a]	0.012	**0.018** [a]	0.012	-0.006	0.009
LEP	**-0.027** [a]	0.008	**-0.031** [a]	0.007	**-0.005** [a]	0.008	**-0.011** [a]	0.007	**-0.027** [a]	0.004
504 plan	**-0.008** [a]	0.008	**-0.007** [a]	0.009	**-0.007** [a]	0.010	**-0.003** [a]	0.010	-0.001	0.011
Economic disadvantage	**-0.067** [a]	0.007	**-0.071** [a]	0.007	**-0.011** [a]	0.008	**-0.016** [a]	0.005	**-0.056** [a]	0.005
ELA	**0.552** [a]	0.002	**0.509** [a]	0.002	**0.318** [a]	0.004	**0.326** [a]	0.003	**0.294** [a]	0.003

Note. LEP = limited English proficiency; ELA = English language arts.
[a] Statistically significant.

Table C10. Model 5C Standardized Regression Coefficients With Factor Scores by Outcome Variables: Interaction Effects

Standard	Overall algebra beta	Overall algebra SE	Quantitative literacy beta	Quantitative literacy SE	Generalized arithmetic beta	Generalized arithmetic SE	Functional thinking beta	Functional thinking SE	Geometry & statistics beta	Geometry & statistics SE
A1b: female	**-0.029** [a]	0.002	**-0.024** [a]	0.002	**-0.028** [a]	0.002	**-0.012** [a]	0.002	**-0.020** [a]	0.002
A1b: Hispanic or Latino	**-0.012** [a]	0.002	**-0.018** [a]	0.003	**0.020** [a]	0.003	0.005	0.004	**-0.020** [a]	0.003
A1b: American Indian or Alaskan Native	-0.001	0.011	**-0.002** [a]	0.011	**0.004** [a]	0.017	0.001	0.017	0.000	0.016
A1b: Asian	**0.006** [a]	0.006	**0.006** [a]	0.006	0.001	0.007	0.003	0.005	**0.013** [a]	0.006
A1b: Black or African American	**-0.006** [a]	0.005	**-0.012** [a]	0.005	**0.016** [a]	0.006	**0.006** [a]	0.006	**-0.008** [a]	0.006
A1b: Native Hawaiian or Other Pacific Islander	**-0.002** [a]	0.011	**-0.002** [a]	0.012	-0.001	0.016	0.000	0.015	-0.002	0.016
A1b: Filipino	**-0.003** [a]	0.006	**-0.003** [a]	0.006	**-0.004** [a]	0.008	**0.004** [a]	0.007	0.000	0.007
A1b: LEP	-0.004	0.003	**-0.011** [a]	0.003	**0.034** [a]	0.003	-0.001	0.004	**-0.011** [a]	0.003
A1b: 501b plan	-0.001	0.007	-0.001	0.007	0.000	0.009	0.000	0.010	-0.001	0.010
A1b: Economic disadvantage	**-0.012** [a]	0.002	**-0.018** [a]	0.002	**0.019** [a]	0.004	**0.006** [a]	0.003	**-0.028** [a]	0.003

Note. LEP = limited English proficiency.
[a] Statistically significant.

Table C11. Model 5D Standardized Regression Coefficients With Factor Scores by Outcome Variables: Main Effects

Standard	Overall algebra beta [a]	Overall algebra SE	Quantitative literacy beta [a]	Quantitative literacy SE	Generalized arithmetic beta [a]	Generalized arithmetic SE	Functional thinking beta [a]	Functional thinking SE	Geometry & statistics beta [b]	Geometry & statistics SE
A3a—CPS whole numbers	**0.435**	0.003	**0.442**	0.003	**0.142**	0.004	**0.175**	0.003	**0.224**	0.003
Female	**-0.050**	0.002	**-0.054**	0.002	**-0.004**	0.002	**0.007**	0.002	**-0.043**	0.002
Hispanic or Latino	**-0.031**	0.006	**-0.040**	0.005	**0.036**	0.007	**0.007**	0.005	**-0.064**	0.006
American Indian or Alaskan Native	**-0.008**	0.013	**-0.007**	0.012	**-0.002**	0.018	**-0.006**	0.016	**-0.005**	0.015
Asian	**0.090**	0.015	**0.084**	0.016	**0.065**	0.012	**0.047**	0.010	**0.031**	0.010
Black or African American	**-0.040**	0.009	**-0.046**	0.008	**0.004**	0.010	**-0.006**	0.009	**-0.034**	0.008
Native Hawaiian or Other Pacific Islander	**-0.001**	0.014	**-0.002**	0.014	0.004	0.018	**0.001**	0.017	**-0.004**	0.015
Filipino	**0.017**	0.011	**0.016**	0.010	**0.014**	0.012	**0.018**	0.011	**-0.009**	0.009
LEP	**-0.038**	0.009	**-0.051**	0.008	**0.033**	0.009	**-0.020**	0.008	**-0.042**	0.004
504 plan	**-0.006**	0.009	**-0.004**	0.009	**-0.006**	0.010	**-0.004**	0.010	-0.001	0.011
Economic disadvantage	**-0.051**	0.007	**-0.059**	0.007	**0.013**	0.008	**-0.001**	0.005	**-0.063**	0.005
ELA	**0.363**	0.002	**0.325**	0.003	**0.220**	0.004	**0.242**	0.003	**0.223**	0.003

Note. CPS = complex problem solving; LEP = limited English proficiency; ELA = English language arts.
[a] All values in the beta column are statistically significant. [b] All values in the Geometry & statistics column are statistically except for the value for 504 plan.

Table C12. Model 5D Standardized Regression Coefficients With Factor Scores by Outcome Variables: Interaction Effects

Standard	Overall algebra beta	Overall algebra SE	Quantitative literacy beta	Quantitative literacy SE	Generalized arithmetic beta	Generalized arithmetic SE	Functional thinking beta	Functional thinking SE	Geometry & statistics beta	Geometry & statistics SE
A3a: female	**-0.034** [a]	0.002	**-0.026** [a]	0.002	**-0.029** [a]	0.002	**-0.017** [a]	0.002	**-0.033** [a]	0.002
A3a: Hispanic or Latino	**-0.007** [a]	0.003	**-0.016** [a]	0.003	**0.042** [a]	0.004	0.006	0.004	**-0.040** [a]	0.003
A3a: American Indian or Alaskan Native	**-0.002** [a]	0.012	**-0.003** [a]	0.012	0.003	0.018	-0.001	0.016	**-0.004** [a]	0.016
A3a: Asian	0.001	0.007	0.001	0.007	-0.001	0.007	**-0.004** [a]	0.006	**0.015** [a]	0.006
A3a: Black or African American	**-0.011** [a]	0.005	**-0.023** [a]	0.005	**0.029** [a]	0.005	**0.009** [a]	0.006	**-0.015** [a]	0.006
A3a: Native Hawaiian or Other Pacific Islander	**-0.001** [a]	0.011	**-0.003** [a]	0.012	0.003	0.016	0.001	0.016	0.000	0.016
A3a: Filipino	-0.003	0.006	-0.003	0.007	**-0.005** [a]	0.008	0.005	0.008	-0.001	0.007
A3a: LEP	**-0.028** [a]	0.004	**-0.044** [a]	0.004	**0.064** [a]	0.004	**-0.018** [a]	0.004	**-0.028** [a]	0.003
A3a: 504 plan	**-0.001** [a]	0.007	-0.001	0.008	0.000	0.011	-0.002	0.011	-0.002	0.010
A3a: Economic disadvantage	**-0.007** [a]	0.003	**-0.020** [a]	0.003	**0.046** [a]	0.004	**0.020** [a]	0.003	**-0.047** [a]	0.003

Note. LEP = limited English proficiency.
[a] Statistically significant.

Table C13. *R*-squared Statistics for Regression Models With Factor Scores by Outcome Variables

Model	Overall algebra	Quantitative literacy	Generalized arithmetic	Functional thinking	Geometry & statistics	ELA
1	0.618	0.564	0.164	0.162	0.156	0.522
2	0.641	0.581	0.176	0.177	0.168	0.659
3	0.656	0.597	0.179	0.175	0.164	0.551
4	0.665	0.603	0.184	0.183	0.171	0.662
5A	0.617	0.558	0.175	0.174	0.169	-
5B	0.581	0.523	0.164	0.167	0.163	-
5C	0.573	0.515	0.161	0.165	0.161	-
5D	0.572	0.513	0.163	0.165	0.161	-

Note. ELA= English language arts.
[a] Statistically significant.

Appendix D. Extended Regression Models

As part of the regression analyses, Models 5A, 5B, 5C, and 5D were fit to examine the main effects, respectively, for A4—Reasoning and Communicating; A3b—Complex Problem Solving with Rational Numbers; A1b—Number Sense with Fractions; and A3a—Complex Problem Solving with Whole Numbers. The following description reiterates the specification for Model 5A.

- Model 5A—Includes main effects for A4 (the strongest predictor of algebra, based on the previous models), ELA score, and a range of demographic variables; additionally, interaction effects between A4 and the demographic variables are specified (the interaction effects are denoted by the colons, e.g., A4: female is the effect for females with a given A4 score)
    - A4, female, Hispanic or Latino, American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islander, Filipino, limited English proficient (LEP), 504 plan, economic disadvantage, ELA, A4: female, A4: Hispanic or Latino, A4: American Indian or Alaska Native, A4: Asian, A4: Black or African, American, A4: Native Hawaiian or Other Pacific Islander, A4: Filipino, A4: LEP, A4: 504 plan, A4: economic disadvantage

While not the primary focus of the regression analyses, a subcomponent of the research question related to demographic interaction effects relates to potential differences between males and females. As such, an additional set of analyses were conducted that directly extend the specifications of these four models. The extended models, 6A, 6B, 6C, and 6D include the main effects and interaction effects for Models 5A through 5D with additional interaction effects between the demographic variables and gender. The specification for Model 6A is shown below.

- Model 6A—A4, female, Hispanic or Latino, American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islander, Filipino, limited English proficient (LEP), 504 Plan, Economic Disadvantage, ELA, A4: female, A4: Hispanic or Latino, A4: American Indian or Alaska Native, A4: Asian, A4: Black or African, American, A4: Native Hawaiian or Other Pacific Islander, A4: Filipino, A4: LEP, A4: 504 plan, A4:

economic disadvantage, female: Hispanic or Latino, female: American Indian or Alaska Native, female: Asian, female: Black or African, American, female: Native Hawaiian or Other Pacific Islander, female: Filipino, female: LEP, female: 504 plan, female: economic disadvantage

As with Models 5A, 5B, 5C, and 5D, five outcome measures were considered: overall algebra, quantitative literacy, generalized arithmetic, functional thinking, and geometry and statistics. The regression coefficients for the models are presented in Tables D1 through D13. The differences in the coefficients for the effects modeled in Models 5A, 5B, 5C, and 5D are very small. Further, the addition effects for the interactions with gender are extremely small. This suggests that with respect to the interactions, there are no practically significant differences between males and females among the various demographic groups.

As a final extension of the regression models, Model 5A was respecified with overall math performance as the primary predictor of interest. The model specification is as follows.

- Model 6D—Math, female, Hispanic or Latino, American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islander, Filipino, limited English proficient (LEP), 504 plan, economic disadvantage, ELA, math: female, math: Hispanic or Latino, math: American Indian or Alaska Native, math: Asian, math: Black or African, American, math: Native Hawaiian or Other Pacific Islander, math: Filipino, math: LEP, math: 504 plan, math: economic disadvantage

The same five outcome measures were considered. The coefficients for this model are shown in Table D14. As expected, the main effect for overall math performance is the strongest predictor, followed by the effect for ELA score. After controlling for these variables, the main effects for the demographic variables and the interaction effects are very small.

Table D1. Model 6A Factor Scores Regression Effects by Student Group With Predictor A4: Main Effects

Standard	Overall algebra beta	Overall algebra SE	Quantitative literacy beta	Quantitative literacy SE	Generalized arithmetic beta	Generalized arithmetic SE	Functional thinking beta	Functional thinking SE	Geometry & statistics beta	Geometry & statistics SE
A4—Reasoning & communicating	**0.432** [a]	0.003	**0.439** [a]	0.003	**0.140** [a]	0.006	**0.172** [a]	0.004	**0.222** [a]	0.004
Female	**-0.057** [a]	0.003	**-0.063** [a]	0.003	**0.001** [a]	0.005	**0.001** [a]	0.004	**-0.041** [a]	0.005
Hispanic or Latino	**-0.034** [a]	0.005	**-0.044** [a]	0.005	**0.037** [a]	0.007	0.004	0.006	**-0.062** [a]	0.006
American Indian or Alaskan Native	**-0.007** [a]	0.016	**-0.007** [a]	0.016	**-0.001** [a]	0.022	-0.003	0.023	-0.005	0.023
Asian	**0.090** [a]	0.014	**0.082** [a]	0.014	**0.072** [a]	0.012	**0.047** [a]	0.011	**0.033** [a]	0.009
Black or African American	**-0.045** [a]	0.009	**-0.051** [a]	0.008	**0.005** [a]	0.012	**-0.010** [a]	0.009	**-0.035** [a]	0.009
Native Hawaiian or Other Pacific Islander	-0.001	0.018	-0.002	0.017	0.004	0.023	0.000	0.024	**-0.005** [a]	0.024
Filipino	**0.017** [a]	0.011	**0.014** [a]	0.010	**0.020** [a]	0.012	**0.018** [a]	0.014	**-0.008** [a]	0.011
LEP	**-0.038** [a]	0.009	**-0.051** [a]	0.008	0.036	0.008	**-0.021** [a]	0.008	**-0.045** [a]	0.005
504 plan	**-0.005** [a]	0.010	**-0.004** [a]	0.010	**-0.005** [a]	0.013	-0.003	0.012	0.000	0.013
Economic disadvantage	**-0.054** [a]	0.006	**-0.062** [a]	0.006	**0.010** [a]	0.008	**-0.004** [a]	0.005	**-0.063** [a]	0.005
ELA	**0.363** [a]	0.002	**0.325** [a]	0.002	**0.220** [a]	0.004	**0.242** [a]	0.003	**0.223** [a]	0.002

Note. LEP = limited English proficiency; ELA = English language arts.

[a] Statistically significant.

Table D2. Model 6A Factor Scores Regression Effects by Student Group With Predictor A4: Interaction Effects

Standard	Overall algebra beta	Overall algebra SE	Quantitative literacy beta	Quantitative literacy SE	Generalized arithmetic beta	Generalized arithmetic SE	Functional thinking beta	Functional thinking SE	Geometry & statistics beta	Geometry & statistics SE
A4: female	**-0.029** [a]	0.002	**-0.022** [a]	0.002	**-0.026** [a]	0.003	**-0.012** [a]	0.002	**-0.030** [a]	0.003
A4: Hispanic or Latino	**-0.007** [a]	0.003	**-0.016** [a]	0.003	**0.042** [a]	0.004	0.007	0.004	**-0.040** [a]	0.003
A4: American Indian or Alaskan Native	-0.002	0.010	**-0.003** [a]	0.011	0.003	0.015	-0.001	0.017	-0.004	0.017
A4: Asian	0.001	0.007	0.001	0.007	-0.001	0.008	-0.004	0.007	**0.015** [a]	0.007
A4: Black or African American	**-0.011** [a]	0.006	**-0.023** [a]	0.006	**0.029** [a]	0.007	**0.009** [a]	0.006	**-0.015** [a]	0.006
A4: Native Hawaiian or Other Pacific Islander	-0.001	0.010	-0.002	0.010	0.003	0.016	0.001	0.017	0.000	0.018
A4: Filipino	**-0.003** [a]	0.006	**-0.003** [a]	0.006	**-0.005** [a]	0.008	**0.005** [a]	0.008	-0.001	0.008
A4: LEP	**-0.028** [a]	0.005	**-0.044** [a]	0.006	**0.063** [a]	0.004	**-0.017** [a]	0.004	**-0.028** [a]	0.003
A4: 504 plan	-0.001	0.007	-0.001	0.007	0.000	0.012	-0.002	0.011	-0.002	0.011
A4: Economic disadvantage	**-0.007** [a]	0.003	**-0.020** [a]	0.003	**0.046** [a]	0.004	**0.019** [a]	0.004	**-0.047** [a]	0.003

Note. LEP = limited English proficiency.
[a] Statistically significant.

Table D3. Model 6A Factor Scores Regression Effects by Student Group With Predictor Female: Interaction Effects

Standard	Overall algebra beta	Overall algebra SE	Quantitative literacy beta	Quantitative literacy SE	Generalized arithmetic beta	Generalized arithmetic SE	Functional thinking beta	Functional thinking SE	Geometry & statistics beta	Geometry & statistics SE
Female: Hispanic or Latino	**0.007** [a]	0.004	**0.007** [a]	0.004	-0.001	0.005	**0.007** [a]	0.006	-0.003	0.005
Female: American Indian or Alaskan Native	-0.001	0.023	0.000	0.023	-0.002	0.029	-0.003	0.032	-0.001	0.032
Female: Asian	0.001	0.005	**0.003** [a]	0.005	**-0.010** [a]	0.008	0.001	0.007	-0.003	0.010
Female: Black or African American	**0.007** [a]	0.007	**0.006** [a]	0.007	-0.001	0.009	**0.006** [a]	0.009	0.001	0.009
Female: Native Hawaiian or Other Pacific Islander	0.000	0.021	0.000	0.021	0.000	0.033	0.001	0.030	0.001	0.030
Female: Filipino	0.000	0.010	0.002	0.010	**-0.009** [a]	0.015	0.000	0.015	-0.002	0.014
Female: LEP	0.000	0.004	0.000	0.004	**-0.005** [a]	0.005	0.002	0.005	**0.005** [a]	0.005
Female: 504 plan	**-0.002** [a]	0.015	-0.002	0.016	-0.001	0.022	-0.001	0.020	0.000	0.020
Female: Economic disadvantage	**0.007** [a]	0.003	**0.007** [a]	0.004	**0.005** [a]	0.005	0.005	0.005	0.002	0.005

Note. LEP = limited English proficiency.
[a] Statistically significant.

Table D4. Model 6B Factor Scores Regression Effects by Student Group With Predictor A3b: Main Effects

Standard	Overall algebra beta	Overall algebra SE	Quantitative literacy beta	Quantitative literacy SE	Generalized arithmetic beta	Generalized arithmetic SE	Functional thinking beta	Functional thinking SE	Geometry & statistics beta	Geometry & statistics SE
A3b—CPS fractions	**0.254** [a]	0.003	**0.264** [a]	0.003	**0.066** [a]	0.005	**0.096** [a]	0.003	**0.142** [a]	0.003
Female	**-0.084** [a]	0.003	**-0.088** [a]	0.004	**-0.017** [a]	0.005	-0.010	0.004	**-0.054** [a]	0.005
Hispanic or Latino	**-0.046** [a]	0.006	**-0.057** [a]	0.006	**0.035** [a]	0.007	0.000	0.006	**-0.070** [a]	0.006
American Indian or Alaskan Native	**-0.008** [a]	0.017	**-0.008** [a]	0.017	**0.001** [a]	0.022	**-0.002** [a]	0.023	**-0.007** [a]	0.022
Asian	**0.106** [a]	0.014	**0.098** [a]	0.015	**0.076** [a]	0.013	**0.053** [a]	0.010	**0.046** [a]	0.011
Black or African American	**-0.057** [a]	0.009	**-0.064** [a]	0.009	**0.004** [a]	0.011	**-0.013** [a]	0.009	**-0.045** [a]	0.009
Native Hawaiian or Other Pacific Islander	**-0.004** [a]	0.018	**-0.005** [a]	0.017	0.003	0.023	-0.002	0.024	**-0.008** [a]	0.025
Filipino	**0.017** [a]	0.012	**0.013** [a]	0.011	**0.019** [a]	0.013	**0.018** [a]	0.014	**-0.007** [a]	0.011
LEP	**-0.040** [a]	0.010	**-0.051** [a]	0.009	**0.031** [a]	0.009	**-0.020** [a]	0.008	**-0.047** [a]	0.005
504 plan	**-0.008** [a]	0.010	**-0.006** [a]	0.011	**-0.007** [a]	0.013	-0.005	0.012	-0.001	0.013
Economic disadvantage	**-0.069** [a]	0.007	**-0.077** [a]	0.007	**0.004** [a]	0.008	**-0.009** [a]	0.005	**-0.072** [a]	0.005
ELA	**0.508** [a]	0.002	**0.467** [a]	0.002	**0.293** [a]	0.004	**0.306** [a]	0.003	**0.278** [a]	0.003

Note. CPS = complex problem solving; LEP = limited English proficiency; ELA = English language arts.

[a] Statistically significant.

Table D5. Model 6B Factor Scores Regression Effects by Student Group With Predictor A3b: Interaction Effects

Standard	Overall algebra beta	Overall algebra SE	Quantitative literacy beta	Quantitative literacy SE	Generalized arithmetic beta	Generalized arithmetic SE	Functional thinking beta	Functional thinking SE	Geometry & statistics beta	Geometry & statistics SE
A3b: Female	**-0.025** [a]	0.001	**-0.019** [a]	0.002	**-0.028** [a]	0.002	**-0.009** [a]	0.002	**-0.026** [a]	0.002
A3b: Hispanic or Latino	**-0.009** [a]	0.003	**-0.017** [a]	0.003	**0.039** [a]	0.004	0.006	0.004	**-0.038** [a]	0.003
A3b: American Indian or Alaskan Native	-0.001	0.011	-0.002	0.012	**0.006** [a]	0.015	0.001	0.017	-0.004	0.016
A3b: Asian	0.002	0.007	0.002	0.007	-0.003	0.008	-0.003	0.006	**0.016** [a]	0.007
A3b: Black or African American	**-0.014** [a]	0.005	**-0.026** [a]	0.006	**0.031** [a]	0.007	**0.010** [a]	0.006	**-0.020** [a]	0.006
A3b: Native Hawaiian or Other Pacific Islander	-0.002	0.011	-0.003	0.012	0.002	0.015	-0.001	0.018	-0.003	0.015
A3b: Filipino	**-0.004** [a]	0.006	**-0.004** [a]	0.006	**-0.006** [a]	0.008	0.004	0.007	0.000	0.008
A3b: LEP	**-0.022** [a]	0.004	**-0.037** [a]	0.005	**0.061** [a]	0.004	**-0.011** [a]	0.004	**-0.027** [a]	0.003
A3b: 504 plan	-0.002	0.007	-0.001	0.007	-0.002	0.012	-0.003	0.011	-0.002	0.011
A3b: Economic disadvantage	-0.005	0.003	**-0.015** [a]	0.003	**0.036** [a]	0.004	**0.018** [a]	0.003	**-0.037** [a]	0.003

Note. LEP = limited English proficiency.
[a] Statistically significant.

Table D6. Model 6B Factor Scores Regression Effects by Student Group With Predictor Female: Interaction Effects

Standard	Overall algebra beta	Overall algebra SE	Quantitative literacy beta	Quantitative literacy SE	Generalized arithmetic beta	Generalized arithmetic SE	Functional thinking beta	Functional thinking SE	Geometry & statistics beta	Geometry & statistics SE
Female: Hispanic or Latino	**0.011** [a]	0.004	**0.011** [a]	0.004	0.001	0.005	**0.008** [a]	0.006	-0.001	0.005
Female: American Indian or Alaskan Native	-0.001	0.023	0.000	0.024	-0.002	0.028	-0.003	0.032	0.000	0.032
Female: Asian	0.000	0.006	**0.003** [a]	0.006	**-0.011** [a]	0.008	0.000	0.007	**-0.004** [a]	0.010
Female: Black or African American	**0.010** [a]	0.008	**0.009** [a]	0.007	0.000	0.009	**0.007** [a]	0.010	**0.003** [a]	0.009
Female: Native Hawaiian or Other Pacific Islander	0.001	0.023	0.001	0.023	0.000	0.034	0.002	0.031	0.001	0.030
Female: Filipino	**0.002** [a]	0.009	**0.004** [a]	0.009	**-0.008** [a]	0.015	0.001	0.015	-0.001	0.014
Female: LEP	**0.004** [a]	0.004	0.002	0.004	**-0.003** [a]	0.005	**0.004** [a]	0.005	**0.008** [a]	0.005
Female: 504 plan	**-0.002** [a]	0.016	-0.002	0.016	-0.001	0.022	-0.001	0.020	0.000	0.020
Female: Economic disadvantage	**0.014** [a]	0.003	**0.013** [a]	0.004	**0.008** [a]	0.005	**0.008** [a]	0.005	**0.006** [a]	0.005

Note. LEP = limited English proficiency.

[a] Statistically significant.

Table D7. Model 6C Factor Scores Regression Effects by Student Group With Predictor A1b: Main Effects

Standard	Overall algebra beta	Overall algebra SE	Quantitative literacy beta	Quantitative literacy SE	Generalized arithmetic beta	Generalized arithmetic SE	Functional thinking beta	Functional thinking SE	Geometry & statistics beta	Geometry & statistics SE
A1b—fractions	**0.214** [a]	0.003	**0.225** [a]	0.003	**0.043** [a]	0.004	**0.073** [a]	0.003	**0.118** [a]	0.003
Female	**-0.061** [a]	0.003	**-0.067** [a]	0.004	**-0.006** [a]	0.004	-0.002	0.004	**-0.034** [a]	0.005
Hispanic or Latino	**-0.042** [a]	0.006	**-0.049** [a]	0.006	0.015	0.007	-0.002	0.006	**-0.050** [a]	0.006
American Indian or Alaskan Native	**-0.007** [a]	0.017	**-0.007** [a]	0.017	**-0.001** [a]	0.021	**-0.003** [a]	0.022	-0.003	0.022
Asian	**0.112** [a]	0.014	**0.103** [a]	0.014	**0.080** [a]	0.012	**0.058** [a]	0.010	**0.039** [a]	0.012
Black or African American	**-0.049** [a]	0.009	**-0.048** [a]	0.008	**-0.013** [a]	0.011	**-0.019** [a]	0.009	**-0.033** [a]	0.009
Native Hawaiian or Other Pacific Islander	-0.003	0.019	-0.003	0.018	0.001	0.023	-0.001	0.024	**-0.007** [a]	0.024
Filipino	**0.020** [a]	0.012	**0.018** [a]	0.011	**0.023** [a]	0.013	**0.017** [a]	0.015	**-0.006** [a]	0.011
LEP	**-0.027** [a]	0.008	**-0.030** [a]	0.007	**-0.003** [a]	0.008	**-0.012** [a]	0.007	**-0.031** [a]	0.005
504 plan	**-0.007** [a]	0.011	**-0.006** [a]	0.011	**-0.006** [a]	0.013	**-0.003** [a]	[a]0.012	0.000	0.013
Economic disadvantage	**-0.071** [a]	0.007	**-0.074** [a]	0.007	**-0.015** [a]	0.008	**-0.020** [a]	0.005	**-0.058** [a]	0.005
ELA	**0.552** [a]	0.002	**0.509** [a]	0.002	**0.318** [a]	0.004	**0.326** [a]	0.003	**0.294** [a]	0.003

Note. LEP = limited English proficiency; ELA = English language arts.
[a] Statistically significant.

Table D8. Model 6C Factor Scores Regression Effects by Student Group With Predictor A1b: Interaction Effects

Standard	Overall algebra beta	Overall algebra SE	Quantitative literacy beta	Quantitative literacy SE	Generalized arithmetic beta	Generalized arithmetic SE	Functional thinking beta	Functional thinking SE	Geometry & statistics beta	Geometry & statistics SE
A1b: Female	**-0.026** [a]	0.002	**-0.022** [a]	0.002	**-0.024** [a]	0.002	**-0.008** [a]	0.002	**-0.018** [a]	0.002
A1b: Hispanic or Latino	**-0.011** [a]	0.002	**-0.017** [a]	0.003	**0.020** [a]	0.003	0.006	0.004	**-0.020** [a]	0.003
A1b: American Indian or Alaskan Native	-0.001	0.011	**-0.002** [a]	0.011	**0.004** [a]	0.017	0.001	0.017	0.000	0.016
A1b: Asian	**0.006** [a]	0.006	**0.006** [a]	0.006	0.000	0.007	0.003	0.005	**0.013** [a]	0.006
A1b: Black or African American	**-0.006** [a]	0.005	**-0.012** [a]	0.005	**0.016** [a]	0.006	**0.006** [a]	0.006	**-0.008** [a]	0.006
A1b: Native Hawaiian or Other Pacific Islander	-0.002	0.011	**-0.002** [a]	0.012	-0.001	0.016	0.001	0.015	-0.002	0.016
A1b: Filipino	**-0.002** [a]	0.006	**-0.003** [a]	0.006	**-0.004** [a]	0.009	**0.004** [a]	0.007	0.000	0.007
A1b: LEP	-0.004	0.003	**-0.011** [a]	0.004	**0.034** [a]	0.003	0.000	0.004	**-0.010** [a]	0.003
A1b: 504 plan	-0.001	0.007	-0.001	0.007	0.000	0.009	0.000	0.010	-0.001	0.010
A1b: Economic disadvantage	**-0.010** [a]	0.002	**-0.017** [a]	0.002	**0.021** [a]	0.004	**0.007** [a]	0.003	**-0.028** [a]	0.003

Note. LEP = limited English proficiency.
[a] Statistically significant.

Table D9. Model 6C Factor Scores Regression Effects by Student Group With Predictor Female: Interaction Effects

Standard	Overall algebra beta	Overall algebra SE	Quantitative literacy beta	Quantitative literacy SE	Generalized arithmetic beta	Generalized arithmetic SE	Functional thinking beta	Functional thinking SE	Geometry & statistics beta	Geometry & statistics SE
Female: Hispanic or Latino	**0.007** [a]	0.004	**0.006** [a]	0.004	0.001	0.005	**0.008** [a]	0.006	-0.004	0.005
Female: American Indian or Alaskan Native	-0.001	0.024	0.000	0.024	-0.001	0.029	-0.003	0.032	0.000	0.032
Female: Asian	0.001	0.005	**0.004** [a]	0.006	**-0.011** [a]	0.008	0.001	0.007	-0.003	0.010
Female: Black or African American	**0.008** [a]	0.008	**0.006** [a]	0.008	0.001	0.009	**0.007** [a]	0.010	0.002	0.009
Female: Native Hawaiian or Other Pacific Islander	0.001	0.023	0.001	0.023	0.000	0.035	0.002	0.031	0.001	0.030
Female: Filipino	**0.002** [a]	0.010	**0.004** [a]	0.010	**-0.009** [a]	0.015	0.002	0.015	-0.001	0.015
Female: LEP	0.000	0.004	-0.002	0.004	**-0.003** [a]	0.005	0.002	0.005	**0.008** [a]	0.005
Female: 504 plan	**-0.002** [a]	0.016	-0.002	0.017	-0.001	0.022	-0.001	0.020	0.000	0.021
Female: Economic disadvantage	**0.009** [a]	0.003	**0.007** [a]	0.003	**0.009** [a]	0.005	**0.007** [a]	0.005	0.003	0.005

Note. LEP = limited English proficiency.
[a] Statistically significant.

Table D10. Model 6D Factor Scores Regression Effects by Student Group With Predictor A3a: Main Effects

Standard	Overall algebra beta	Overall algebra SE	Quantitative literacy beta	Quantitative literacy SE	Generalized arithmetic beta	Generalized arithmetic SE	Functional thinking beta	Functional thinking SE	Geometry & statistics beta	Geometry & statistics SE
A3a—CPS whole numbers	**0.227** [a]	0.003	**0.233** [a]	0.003	**0.053** [a]	0.004	**0.082** [a]	0.003	**0.129** [a]	0.003
Female	**-0.081** [a]	0.003	**-0.087** [a]	0.004	-0.009	0.005	-0.011	0.004	**-0.049** [a]	0.005
Hispanic or Latino	**-0.051** [a]	0.006	**-0.059** [a]	0.006	0.022	0.007	**-0.005** [a]	0.006	**-0.060** [a]	0.006
American Indian or Alaskan Native	**-0.009** [a]	0.017	**-0.009** [a]	0.017	**-0.001** [a]	0.022	**-0.003** [a]	0.023	**-0.007** [a]	0.023
Asian	**0.111** [a]	0.015	**0.103** [a]	0.016	**0.079** [a]	0.013	**0.055** [a]	0.011	**0.044** [a]	0.012
Black or African American	**-0.060** [a]	0.009	**-0.062** [a]	0.009	**-0.010** [a]	0.011	**-0.018** [a]	0.009	**-0.041** [a]	0.009
Native Hawaiian or Other Pacific Islander	**-0.004** [a]	0.019	**-0.005** [a]	0.018	0.003	0.023	-0.003	0.024	**-0.007** [a]	0.025
Filipino	**0.018** [a]	0.012	**0.016** [a]	0.011	**0.022** [a]	0.012	**0.017** [a]	0.014	**-0.007** [a]	0.011
LEP	**-0.062** [a]	0.009	**-0.067** [a]	0.009	**-0.012** [a]	0.009	**-0.034** [a]	0.008	**-0.045** [a]	0.005
504 plan	**-0.008** [a]	0.011	**-0.007** [a]	0.011	**-0.006** [a]	0.013	**-0.003** [a]	0.012	-0.001	0.013
Economic disadvantage	**-0.078** [a]	0.008	**-0.085** [a]	0.007	**-0.005** [a]	0.008	**-0.016** [a]	0.005	**-0.070** [a]	0.005
ELA	**0.536** [a]	0.002	**0.497** [a]	0.003	**0.302** [a]	0.004	**0.320** [a]	0.003	**0.290** [a]	0.003

Note. CPS = complex problem solving; LEP = limited English proficiency; ELA = English language arts.
[a] Statistically significant.

Table D11. Model 6D Factor Scores Regression Effects by Student Group With Predictor A3a: Interaction Effects

Standard	Overall algebra beta	Overall algebra SE	Quantitative literacy beta	Quantitative literacy SE	Generalized arithmetic beta	Generalized arithmetic SE	Functional thinking beta	Functional thinking SE	Geometry & statistics beta	Geometry & statistics SE
A3a: Female	**-0.022** [a]	0.002	**-0.017** [a]	0.002	**-0.017** [a]	0.002	**-0.012** [a]	0.002	**-0.023** [a]	0.002
A3a: Hispanic or Latino	**-0.009** [a]	0.003	**-0.016** [a]	0.003	**0.030** [a]	0.004	0.003	0.004	**-0.026** [a]	0.003
A3a: American Indian or Alaskan Native	**-0.003** [a]	0.012	**-0.004** [a]	0.012	0.004	0.018	0.000	0.016	**-0.004** [a]	0.016
A3a: Asian	0.002	0.007	0.003	0.007	-0.003	0.007	**-0.004** [a]	0.006	**0.014** [a]	0.006
A3a: Black or African American	**-0.015** [a]	0.005	**-0.024** [a]	0.005	**0.017** [a]	0.005	**0.005** [a]	0.006	**-0.016** [a]	0.006
A3a: Native Hawaiian or Other Pacific Islander	**-0.002** [a]	0.011	**-0.002** [a]	0.012	0.002	0.016	-0.002	0.016	-0.001	0.016
A3a: Filipino	-0.001	0.006	-0.001	0.007	**-0.004** [a]	0.008	0.002	0.008	0.000	0.007
A3a: LEP	**-0.049** [a]	0.004	**-0.058** [a]	0.004	**0.017** [a]	0.004	**-0.029** [a]	0.004	**-0.028** [a]	0.003
A3a: 504 plan	**-0.002** [a]	0.007	-0.002	0.007	0.000	0.011	-0.001	0.011	-0.002	0.010
A3a: Economic disadvantage	**-0.015** [a]	0.003	**-0.026** [a]	0.003	**0.033** [a]	0.004	**0.012** [a]	0.003	**-0.040** [a]	0.003

Note. LEP = limited English proficiency.
[a] Statistically significant.

Table D12. Model 6D Factor Scores Regression Effects by Student Group With Predictor Female: Interaction Effects

Standard	Overall algebra beta	Overall algebra SE	Quantitative literacy beta	Quantitative literacy SE	Generalized arithmetic beta	Generalized arithmetic SE	Functional thinking beta	Functional thinking SE	Geometry & statistics beta	Geometry & statistics SE
Female: Hispanic or Latino	**0.011** [a]	0.004	**0.011** [a]	0.004	0.002	0.006	**0.008** [a]	0.006	-0.001	0.005
Female: American Indian or Alaskan Native	-0.001	0.024	0.000	0.025	-0.002	0.029	-0.003	0.032	0.000	0.032
Female: Asian	0.000	0.006	**0.003** [a]	0.006	**-0.012** [a]	0.008	0.001	0.007	-0.004	0.010
Female: Black or African American	**0.010** [a]	0.008	**0.009** [a]	0.008	0.001	0.009	**0.007** [a]	0.010	**0.003** [a]	0.009
Female: Native Hawaiian or Other Pacific Islander	0.001	0.023	0.001	0.023	0.001	0.034	0.002	0.031	0.001	0.030
Female: Filipino	**0.002** [a]	0.010	**0.004** [a]	0.010	**-0.009** [a]	0.015	0.001	0.015	-0.001	0.014
Female: LEP	**0.004** [a]	0.004	**0.003** [a]	0.004	-0.002	0.005	**0.003** [a]	0.005	**0.008** [a]	0.005
Female: 504 plan	**-0.002** [a]	0.017	**-0.002** [a]	0.017	-0.001	0.022	-0.001	0.020	0.000	0.020
Female: Economic disadvantage	**0.012** [a]	0.004	**0.011** [a]	0.004	**0.009** [a]	0.005	**0.006** [a]	0.005	**0.005** [a]	0.005

Note. LEP = limited English proficiency.
[a] Statistically significant.

Table D13. Model 6E Factor Scores Regression Effects by Student Group With Predictor Overall Math: Main Effects

Standard	Overall algebra beta	Overall algebra SE	Quantitative literacy beta	Quantitative literacy SE	Generalized arithmetic beta	Generalized arithmetic SE	Functional thinking beta	Functional thinking SE	Geometry & statistics beta	Geometry & statistics SE
Math	**0.637** [a]	0.001	**0.640** [a]	0.001	**0.238** [a]	0.001	**0.264** [a]	0.001	**0.306** [a]	0.000
Female	**0.002** [a]	0.001	**-0.007** [a]	0.001	**0.028** [a]	0.001	**0.032** [a]	0.001	**-0.013** [a]	0.000
Hispanic or Latino	**-0.015** [a]	0.001	**-0.022** [a]	0.001	**0.023** [a]	0.001	**0.010** [a]	0.001	**-0.041** [a]	0.001
American Indian or Alaskan Native	**-0.006** [a]	0.008	**-0.005** [a]	0.009	**-0.003** [a]	0.008	**-0.004** [a]	0.005	**-0.003** [a]	0.003
Asian	**0.071** [a]	0.002	**0.064** [a]	0.002	**0.057** [a]	0.002	**0.042** [a]	0.001	**0.009** [a]	0.001
Black or African American	**-0.028** [a]	0.003	**-0.029** [a]	0.003	**-0.005** [a]	0.003	**-0.006** [a]	0.002	**-0.025** [a]	0.001
Native Hawaiian or Other Pacific Islander	0.000	0.008	0.000	0.008	**0.003** [a]	0.008	0.000	0.005	**-0.004** [a]	0.003
Filipino	**0.018** [a]	0.003	**0.017** [a]	0.004	**0.018** [a]	0.003	**0.014** [a]	0.002	**-0.008** [a]	0.001
LEP	**-0.032** [a]	0.002	**-0.040** [a]	0.002	**0.005** [a]	0.002	**-0.012** [a]	0.001	**-0.033** [a]	0.001
504 plan	**-0.005** [a]	0.005	**-0.004** [a]	0.005	**-0.006** [a]	0.005	**-0.002** [a]	0.003	0.001	0.002
Economic disadvantage	**-0.034** [a]	0.001	**-0.037** [a]	0.001	-0.002	0.001	**-0.004** [a]	0.001	**-0.037** [a]	0.001
ELA	**0.184** [a]	0.001	**0.152** [a]	0.001	**0.126** [a]	0.001	**0.157** [a]	0.001	**0.154** [a]	0.000

Note. LEP = limited English proficiency; ELA = English language arts.
[a] Statistically significant.

Table D14. Model 6E Factor Scores Regression Effects by Student Group With Predictor Overall Math: Interaction Effects

Standard	Overall algebra beta	Overall algebra SE	Quantitative literacy beta	Quantitative literacy SE	Generalized arithmetic beta	Generalized arithmetic SE	Functional thinking beta	Functional thinking SE	Geometry & statistics beta	Geometry & statistics SE
Math: Female	**-0.002** [a]	0.001	**0.004** [a]	0.001	**-0.012** [a]	0.001	-0.002	0.001	**-0.017** [a]	0.000
Math: Hispanic or Latino	**-0.011** [a]	0.001	**-0.019** [a]	0.001	**0.035** [a]	0.001	**0.006** [a]	0.001	**-0.032** [a]	0.001
Math: American Indian or Alaskan Native	**-0.004** [a]	0.006	**-0.005** [a]	0.007	**0.003** [a]	0.006	0.000	0.004	**-0.003** [a]	0.003
Math: Asian	-0.001	0.002	-0.001	0.002	0.001	0.002	**-0.004** [a]	0.001	**0.017** [a]	0.001
Math: Black or African American	**-0.019** [a]	0.002	**-0.028** [a]	0.002	**0.017** [a]	0.002	**0.003** [a]	0.002	**-0.016** [a]	0.001
Math: Native Hawaiian or Other Pacific Islander	**-0.002** [a]	0.007	**-0.003** [a]	0.007	0.001	0.007	0.000	0.005	-0.001	0.003
Math: Filipino	0.010	0.003	0.000	0.003	**-0.004** [a]	0.003	**0.004** [a]	0.002	0.001	0.001
Math: LEP	**-0.031** [a]	0.001	**-0.044** [a]	0.001	**0.038** [a]	0.001	**-0.014** [a]	0.001	**-0.022** [a]	0.001
Math: 504 plan	**-0.002** [a]	0.004	**-0.002** [a]	0.004	-0.001	0.004	-0.002	0.003	-0.002	0.002
Math: Economic disadvantage	**-0.009** [a]	0.001	**-0.020** [a]	0.001	**0.036** [a]	0.001	**0.017** [a]	0.001	**-0.041** [a]	0.000

Note. LEP = limited English proficiency.
[a] Statistically significant.