Florida State University Libraries

2021

Elementary Mathematics Student Assessment: Measuring the Performance of Grade K, 1, 2, and 3 Students Number (Whole numbers and Fractions), Operations, and Algebraic Thinking in Spring 2019

Robert C. Schoen, Charity Buntin, Ahmet Guven and Xiaotong Yang

Elementary Mathematics Student Assessment

Measuring the Performance of Grade K, 1, 2, and 3
Students in Number (Whole numbers and Fractions),
Operations, and Algebraic Thinking in Spring 2019

Robert C. Schoen
Charity Buntin
Ahmet Guven
Xiaotong Yang

MARCH 2021

Research Report No. 2021-03

SECURE VERSION

The research and development reported here were supported by the Institute of Education Sciences, U.S. Department of Education, through Award Number R305A170463 to Florida State University. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

Suggested citation: Schoen, R. C., Buntin, C., Guven, A., & Yang, X. (2021). *Elementary mathematics student assessment: Measuring the performance of grade K, 1, 2, and 3 students in number (whole numbers and fractions), operations, and algebraic thinking in spring 2019*. (Research Report No. 2021-03). Florida State University. DOI:10.33009/fsu-1622059548

Copyright 2020, Florida State University. All rights reserved. Requests for permission to use this test in whole or in part should be directed to Robert Schoen, rschoen@lsi.fsu.edu, FSU Learning Systems Institute, 4600 University Center C, Tallahassee, FL, 32306

Detailed information about items is not included in this report. This information was removed so that the psychometric report could be released without compromising test security. Requests to view the full report should be directed to Robert Schoen (rschoen@lsi.fsu.edu).

Elementary Mathematics Student Assessment

Measuring the Performance of Grade K, 1, 2, and 3 Students in Number (Whole Numbers and Fractions), Operations, and Algebraic Thinking in Spring 2019

Research Report No. 2021-03

Robert C. Schoen

Charity Buntin

Ahmet Guven

Xiaotong Yang

March 2021

Florida State University

Florida Center for Research in Science, Technology, Engineering, and Mathematics (FCR-STEM)
Learning Systems Institute
Florida State University
Tallahassee, FL 32306
(850) 644-2570

Acknowledgements

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant Award No. R305A170463 to Florida State University. The opinions expressed are those of the authors and do not represent views of the institute or the U.S. Department of Education.

The successful development and implementation of this assessment involved many experts in mathematics education and many more students. Some of the key people involved in the development of the assessment are listed here along with their roles in the endeavor.

Robert Schoen designed the content and format of the test, coordinated the feasibility study, created the scoring criteria, interpreted the results, and coordinated the writing of this report. Charity Buntin coordinated the review of alignment of items with the Mathematics Florida Standards, assisted with the development and production of the test, coordinated the dissemination and collection of the tests and corresponding consent forms, managed the report-writing process, and assisted with editing the style and format of the report. Ahmet Guven and Xiaotong Yang performed the data analysis for initial item review, scoring, dimensionality, classical test-theory models, and item-response-theory models and shared the writing of the corresponding sections of the report.

Many others were involved in the work that was required to complete this task. In the role of principal investigator for the overall research study, Colleen Ganley managed the application and administration of the protocol approved by the Institutional Review Board at Florida State University and recruited districts, schools, and teachers to participate in the study. Claire Riddell assisted with the development and production of the test and the scoring criteria. Kristy Farina designed and managed the data-entry system, trained data-entry personnel on the system, verified the accuracy of the data, and assisted with description of the data-entry process and sample descriptives for the present report. Caleb Iszler, Victoria Krentz, Rachel Gaitan, Sasha Campbell, and Alexander Dabb entered the data for the tests. Robert Schoen, Charity Buntin, Arias Ware, Anne Kelley, Gizem Solmaz, Caleb Iszler, Victoria Krentz, Rachel Gaitan, Sasha Campbell, and Alexander Dabbs served on the adjudication committee. Insu Paek provided advice throughout the data analysis and interpretation process. Anne Thistle provided valuable assistance with copy editing.

We are especially grateful for the financial support from the Institute of Education Sciences at the U.S. Department of Education and for the students, parents, principals, district leaders, and teachers who agreed to participate in the study.

Table of Contents

Acknowledgements	iv
Executive Summary	xvi
Purpose	xvi
Content	xvi
Sample and Setting	xvi
Test Specifications and Administration	xvi
Data Entry and Scoring	xvii
Reliability	xvii
1. Introduction and Overview	1
1.1. Test Overview	1
1.1.1. Counting and Cardinality	2
1.1.2. Word Problems	2
1.1.3. Number Relations Section	3
1.1.4. Fractions Section	4
1.1.5. Basic Number Facts Section	5
1.1.6. Computation Section	5
1.1.7. Detailed Test Blueprint	6
1.2. Description of the Sample and Setting	9
2. Test Development, Scoring, and Data-Entry Procedures	10
2.1. Content	10
2.2. Instrument Development Process	10
2.3. Test Design and Assembly	11
2.4. Test Production	13
2.5. Administration Protocol for the Spring 2019 EMSA K–3 Test	14
2.6. Data Entry and Verification Procedures	16
2.7. Item-scoring Procedures	16
3. Data and Scoring	17
3.1. Missing Response Data	17
3.2. Initial Item Analysis and Item Removal	19

4. Dimensionality Analysis	24
4.1. Parallel Analysis	24
4.2. Confirmatory Factor Analysis	27
5. Classical Test Theory Analyses	32
5.1. Distribution of the Observed Test Score	32
5.2. Item Difficulty and Discrimination	35
5.3. Coefficients α and Standard Error of Measurement	39
6. Item Response Theory (IRT) Analyses	40
6.1. Model Description	40
6.2. Item Difficulty and Discrimination	41
6.3. Test Information Function	53
6.4. Latent Trait Estimation	57
References	66

List of Appendixes

Appendix A. Grade K Test	69
Appendix B. Grade 1 Test	86
Appendix C. Grade 2 Test	105
Appendix D. Grade 3 Test	126
Appendix E. Grade K Administration Guide	148
Appendix F. Grade 1 Administration Guide	168
Appendix G. Grade 2 Administration Guide	189
Appendix H. Grade 3 Administration Guide	211
Appendix I. Scoring Criteria	232
Appendix J. Within-grade IRT Estimates	242
J.1. Within-grade Item Parameter Estimates	242
J.2. Within-grade Test Information and Estimated Latent Trait	254
Appendix K. EMSA Spring 2018 Sample	266
K.1. Item Response Theory Model	270
K.2. Person Scoring and Test Information Function	270
Appendix L. Most Common Incorrect Responses for Each Item	287

List of Tables

Table 1.1. Final Blueprint for the Spring 2019 K–3 EMSA Test	1
Table 1.2. Items in the Counting and Cardinality Section	2
Table 1.3. Items in the Word Problems Section	2
Table 1.4. Items in the Number Relations Section	3
Table 1.5. Items in the Fractions Section	4
Table 1.6. Items in the Basic Number Facts Section	5
Table 1.7. Items in the Computation Section	6
Table 1.8. Detailed Test Blueprint for the Spring 2019 K–3 EMSA	7
Table 1.9. Demographic Characteristics of the Students in the Spring 2019 Field Test of the K–3 EMSA Tests	9
Table 2.1. Number of Constructed- and Selected-Responses on Spring 2019 K-3 EMSA	12
Table 2.2. Number of Times the Correct Answer Appears in Each Position	13
Table 3.1. Missing Response Frequency in Kindergarten Sample	17
Table 3.3. Missing Response Frequency in Grade 2 Sample	18
Table 3.4. Missing Response Frequency in Grade 3 Sample	19
Table 3.5. Item Indexing and Scoring for Both Test-Form and Final-Scale Formats in Kindergarten	20
Table 3.6. Item Indexing and Scoring for Both Test-Form and Final-Scale Formats in Grade 1	21
Table 3.7. Item Indexing and Scoring for Both Test-Form and Final-Scale Formats in Grade 2	22
Table 3.8. Item Indexing and Scoring for both Test-Form and Final-Scale Formats in Grade 3	23
Table 4.1. Model-fit Statistics of the Test in Each Grade Level	27
Table 4.2. Factor Loadings from Kindergarten Test confirmatory factor analysis (CFA)	28
Table 4.3. Factor Loadings from Grade-1 Test CFA	29
Table 4.4. Factor Loadings from Grade-2 Test CFA	30
Table 4.5. Factor Loadings from Grade-3 Test CFA	31
Table 5.1. Item Difficulty and Discrimination from classical test theory (CTT) Analyses of Kindergarten Test	36
Table 5.2. Item Difficulty and Discrimination from CTT Analyses of Grade-1 Test	36
Table 5.3. Item Difficulty and Discrimination from CTT Analyses of Grade-2 Test	37
Table 5.4. Item Difficulty and Discrimination from CTT Analyses of Grade-3 Test	38

Table 5.5. Descriptive Statistics of CTT-based Item Parameter Estimates of the Test at Each Grade Level	39
Table 5.6. Coefficient α of the Test at Each Grade Level	39
Table 6.1. The Number of Items Common to Grade Levels.....	40
Table 6.2. Descriptive Statistics of Discrimination Estimates and Difficulty Estimates of Each Item in Each Grade Level Test	41
Table 6.3. Parameter Estimates and Standard Errors for Final-Scale Items of Kindergarten Test Modeled by two-parameter logistic (2PL) with fixed item parameter calibration (FIPC) Approach.....	42
Table 6.4. Parameter Estimates and Standard Errors for Final-Scale Items of Grade-1 Test Modeled by 2PL.	43
Table 6.5. Parameter Estimates and Standard Errors for Final-Scale Items of Grade-2 Test Modeled by 2PL with FIPC Approach.....	44
Table 6.6. Parameter Estimates and Standard Errors for Final-Scale Items of Grade-3 Test Modeled by 2PL with FIPC Approach.....	45
Table I.1. Grade K Scoring Key	232
Table I.2. Grade 1 Scoring Key	234
Table I.3. Grade 2 Scoring Key	236
Table J.1. Descriptive Statistics of Discrimination Estimates and Difficulty Estimates of Each Item in Each Grade-Level Test.....	242
Table J.2. Parameter Estimates and Standard Errors for Final-Scale Items of Kindergarten Test Modeled with 2PL	242
Table J.3. Parameter Estimates and Standard Errors for Final-Scale Items of Grade 1 Test Modeled with 2PL	243
Table J.4. Parameter Estimates and Standard Errors for Final-Scale Items of Grade-2 Test Modeled with 2PL	244
Table J.5. Parameter Estimates and Standard Errors for Final-Scale Items of Grade-3 Test Modeled with 2PL	245
Table K.1. Common-item match for kindergarten tests	271
Table K.2. Common-item match for grade-1 tests	272
Table K.3. Common-item match for grade-2 tests	273
Table K.4. Common-item match for grade-3 tests	274
Table L.1. Percentage of Grade K Student Responses by Item	287
Table L.2. Percentage of Grade 1 Student Responses by Item	288
Table L.3. Percentage of Grade 2 Student Responses by Item	289

Table L.4. Percentage of Grade 3 Student Responses by Item 290

List of Figures

Figure 2.1. One of the images used in place of page numbers.....	13
Figure 4.1. Parallel analysis scree plot of kindergarten test.	24
Figure 4.2. Parallel analysis scree plot of grade-1 test.	25
Figure 4.3. Parallel analysis scree plot of grade-2 test.	26
Figure 4.4. Parallel analysis scree plot of grade-3 test.	27
Figure 5.1. Bar graph depicting the distribution of the observed test scores for the kindergarten test...	32
Figure 5.2. Bar graph depicting the distribution of the observed test scores for the grade-1 test.	33
Figure 5.3. Bar graph depicting the distribution of the observed test scores for the grade-2 test.	34
Figure 5.4. Bar graph depicting the distribution of the observed test scores for the grade-3 test.	35
Figure 6.1. Item discrimination estimate (a) of each final-scale item of kindergarten test.....	46
Figure 6.2. Item difficulty estimate (b) of each final-scale item of kindergarten test.	47
Figure 6.3. Item discrimination estimate (a) of each final-scale item of grade-1 test.....	48
Figure 6.4. Item difficulty estimate (b) of each final-scale item of grade-1 test.....	49
Figure 6.5. Item discrimination estimate (a) of each final-scale item of grade-2 test.....	50
Figure 6.6. Item difficulty estimate (b) of each final-scale item of grade-2 test.....	51
Figure 6.7. Item discrimination estimate (a) of each final-scale item of grade-3 test.....	52
Figure 6.8. Item difficulty estimate (b) of each final-scale item of grade-3 test.....	53
Figure 6.9. Test information curve and conditional standard error of measurement (CSEM) for the final scale format of kindergarten test.	54
Figure 6.10. Test information curve and CSEM for the final scale format of grade-1 test.....	55
Figure 6.11. Test information curve and CSEM for the final scale format of grade-2 test.	56
Figure 6.12. Test information curve and CSEM for the final scale format of grade-3 test.	57
Figure 6.13. Maximum likelihood estimate (MLE)–based latent estimates of kindergarten students.....	58
Figure 6.14. Expected a posteriori (EAP)–based latent estimates of kindergarten students.	59
Figure 6.15. MLE-based latent estimates of grade-1 students.	60
Figure 6.16. EAP-based latent estimates of grade-1 students.....	61
Figure 6.17. MLE-based latent estimates of grade-2 students.	62
Figure 6.18. EAP-based latent estimates of grade-2 students.....	63
Figure 6.19. MLE-based latent estimates of grade-3 students.	64

Figure 6.20. EAP-based latent estimates of grade-3 students.....	65
Figure J.1. Within-grade item discrimination estimates for kindergarten test.....	246
Figure J.2. Within-grade item difficulty estimates for kindergarten test.	247
Figure J.3. Within-grade item discrimination estimates for grade-1 test.	248
Figure J.4. Within-grade item difficulty estimates for grade-1 test.	249
Figure J.5. Within-grade item discrimination estimates for grade-2 test.	250
Figure J.6. Within-grade item difficulty estimates for grade-2 test.	251
Figure J.7. Within-grade item discrimination estimates for grade-3 test.	252
Figure J.8. Within-grade item difficulty estimates for grade-3 test.	253
Figure J.9. Within-grade test information curve and conditional standard error of measurement (CSEM) for the final scale format of kindergarten test.	254
Figure J.10. Within-grade test information curve and CSEM for the final scale format of grade-1 test (base scale).	255
Figure J.11. Within-grade test information curve and CSEM for the final scale format of grade-2 test.	256
Figure J.12. Within-grade test information curve and CSEM for the final scale format of grade-3 test.	257
Figure J.13. Within-grade maximum likelihood estimation (MLE)—based latent estimates for kindergarten students.	258
Figure J.14. Within-grade expected a posteriori—based latent estimates of kindergarten students.	259
Figure J.15. Within-grade MLE-based latent estimates of grade-1 students.....	260
Figure J.16. Within-grade EAP-based latent estimates of grade-1 students.	261
Figure J.17. Within-grade MLE-based latent estimates of grade-2 students.....	262
Figure J.18. Within-grade EAP-based latent estimates of grade-2 students.	263
Figure J.19. Within-grade MLE-based latent estimates of grade-3 students.....	264
Figure J.20. Within-grade EAP-based latent estimates of grade-3 students.	265
Figure K.1. The distribution of observed test scores for the kindergarten test.....	267
Figure K.2. Distribution of observed test scores for the grade-1 test.....	268
Figure K.3. Distribution of observed test scores for the grade-2 test.....	269
Figure K.4. Distribution of observed test scores for the grade-3 test.....	270
Figure K.5. Test information curve and CSEM for the final scale format of kindergarten test.....	275
Figure K.6. Test information curve and CSEM for the final scale format of grade-1 test.	276
Figure K.7. Test information curve and CSEM for the final scale format of grade-2 test.	277

Figure K.8. Test information curve and CSEM for the final scale format of grade-3 test.	278
Figure K.9. Maximum-likelihood estimation (MLE)–based latent estimates for kindergarten students.	279
Figure K.10. Expected a posteriori (EAP)–based latent estimates for kindergarten students.....	280
Figure K.11. MLE-based latent estimates for grade-1 students.....	281
Figure K.12. EAP-based latent estimates for grade-1 students.	282
Figure K.13. MLE-based latent estimates for grade-2 students.....	283
Figure K.14. EAP-based latent estimates for grade-2 students.	284
Figure K.15. MLE-based latent estimates for grade-3 students.....	285
Figure K.16. EAP-based latent estimates for grade-3 students.	286

List of Equations

Equation 1. Standard error of measurement (1)	39
Equation 2. Two-parameter logistic (2PL) item response theory (IRT) model (2)	40
Equation 3. Conditional standard error of measurement (CSEM)(3)	53

List of Abbreviations

2PL	Two-parameter logistic
CCSS-M	Common Core State Standards for Mathematics
CFA	Confirmatory Factor Analysis
CFI	Comparative Fit Index
CR	Constructed Response
CSEM	Conditional Standard Error of Measurement
CTT	Classical Test Theory
DNS	Did Not Solve
EAP	Expected A Posteriori
ELL	English Language Learner
EMSA	Elementary Mathematics Student Assessment
FIPC	Fixed Item Parameter Calibration
FSA	Florida Standards Assessment
IEP	Individualized Education Program
IRT	Item Response Theory
JCU	Join Change Unknown
MD	Measurement Division
MG	Multiplication Grouping
MLE	Maximum Likelihood Estimation
NR	No Response
PA	Parallel Analysis
RMSEA	Root Mean Square Error of Approximation
SEM	Standard Error of Measurement
SR	Selected Response
SRMR	Standardized Root Mean Residual
SRU	Separate Result Unknown
SWD	Students with Disabilities
UI	Unclear Intent

Executive Summary

The present report describes the Elementary Mathematics Student Assessment (EMSA) as it was used with grade K, 1, 2, and 3 students in spring 2019. Although several versions of EMSA are available, each designed for different purposes, we will refer to the set of forms discussed here as the Spring 2019 K–3 EMSA.

The Spring 2019 K–3 EMSA was designed to serve as a mathematics achievement test administered to students at the end of the school year. It contained items involving number (whole numbers and fractions), word problems, basic addition and subtraction facts, addition and subtraction involving multidigit numbers, and the meaning of the equals sign in mathematics.

Purpose

The intended use of the Spring 2019 K–3 EMSA test was to measure student mathematics performance in the area of number, operations, and algebraic thinking—the mainstay of the elementary mathematics curriculum—for an observational study examining the effects of teacher math anxiety and student math anxiety on student mathematics performance. The purpose of the current report is to serve as a reference document that describes the content of the test, the development process, the administration protocol, the scoring procedures, and the process we used to create the scores for the Spring 2019 K–3 EMSA.

Content

The Spring 2019 K–3 EMSA test was designed to align with the core content in the number and operations domains in the Common Core State Standards for Mathematics (CCSS-M; NGACBP & CCSSO, 2010) in grades K, 1, 2, and 3. In a few cases, the content of the items goes beyond strictly interpreted content limits implied by the CCSS-M. For example, some equals-sign items involve more than three quantities (e.g., $a + b = c + d$) in grade 1.

Sample and Setting

The 2019 Spring EMSA tests were administered to 2556 participating grade K, 1, 2, and 3 students in 29 schools located in four public school districts in Florida during spring 2019. The sample included 573 grade K, 638 grade 1, 655 grade 2, and 690 grade 3 students.

Test Specifications and Administration

The Spring 2019 K–3 EMSA included selected-response and constructed-response test items at each grade level. On selected-response items, students were asked to mark their answer choices by filling in the bubble beneath the answer choice they thought was correct. Selected-response options were based on common responses students at these grade levels have provided in previous versions of these items presented in a constructed-response format in cognitive interviews and field-tests of paper-pencil tests. On constructed-response items, students were asked to write their answer choices in the box provided. The grade K test form consisted of 17 items, the grade 1 test 22 items, the grade 2 test 25 items, and the grade 3 test 28 items. After a review of item statistics on the basis of classical test theory, item-response theory model calibration, and vertical scaling, the final scale for the grades K, 1, 2, and 3 tests drew on 15, 18, 20, and 24 items, respectively. A test form for grade 4 was created and used, but the achieved sample size was too small to support data modeling.

Data Entry and Scoring

Research assistants performed data entry by typing data from each page of every student test into a database using FileMaker Pro software (Version 16). Student responses were entered as a code for selected response items, or exactly as written for constructed response items. A double-entry procedure was performed with a random sample of 11% of the tests. The scored first and second entries matched on more than 99.8% of the item-level responses that were entered.

Although each item on the Spring 2019 EMSA was designed to have a unique, correct solution, students could generate equivalent responses (e.g., 5, 4 + 1). To accommodate this possibility, an adjudication committee met to review the set of all responses to each item and determine the set of correct answers.

Overall test scores were determined by means of a two-parameter logistic model based on item-response theory. The vertical equating (or scaling) was carried out by Fixed Item Parameter Calibration approach under the common item nonequivalent group design. To place the scales on a common metric, grade 1 was chosen as a base scale (Kolen & Brennan, 2014).

Reliability

After confirming a single, dominant factor in each grade-level test, coefficient alpha (Cronbach, 1951) was calculated as the estimate of reliability coefficient. The alpha coefficients for the grades K–3 tests were found to be 0.77, 0.79, 0.82, and 0.87, respectively. For researchers interested in using students' latent trait scores in secondary analysis, the marginal reliability estimates of the grades K–3 tests were 0.78, 0.80, 0.83, and 0.88, respectively. These marginal reliability estimates were calculated for the scales that were put onto a common metric (with grade 1 as baseline grade).

1. Introduction and Overview

The Spring 2019 K–3 EMSA test forms represent an extension of our work in the development and implementation of previously used EMSA tests (Schoen et al., 2016a, 2016b, 2017, 2021a) and the spring 2014 and spring 2015 Mathematics Performance and Cognition (MPAC) Interviews (Schoen et al., 2016c, 2016d). The Spring 2019 K–3 EMSA contained four test forms, corresponding to grades K, 1, 2, and 3. Like the Fall 2018 K–3 EMSA, it was scored on a single scale, vertically scaled across grades, but the tests were not equated across seasons (i.e., fall, spring). The vertically scaled scores were created to support direct comparison across the four grade levels.

The Spring 2019 K–3 EMSA was designed to be administered to students at the end of the school year to measure students' ability to solve problems involving number, operations, and equality. The test does not directly attempt to measure mathematics knowledge or ability in other domains such as geometry, measurement, probability, or data analysis. The set of Spring 2019 EMSA test forms included a grade 4 form, which was designed to be vertically scaled with the K–3 forms, but the achieved sample of grade-4 students during this administration was too small to be used in the analytic sample, so the grade-4 test has been omitted from the current report.

The current report focuses on the content of the test, the development process, the administration protocol, the scoring procedures, and the process we used to create the final scale for the Spring 2019 K–3 EMSA. Its purpose is to serve as a reference document that describes available evidence to support the substantive and structural validity (Flake, Pek, & Hehman, 2017) and the process we used to create the final scale. Although these elements may provide valuable information to other researchers, they also serve as a reference upon which we can base continual future improvement of our design and field-testing of assessment instruments.

1.1. Test Overview

Table 1.1 provides an overall blueprint for each of the four tests.

Table 1.1. Final Blueprint for the Spring 2019 K–3 EMSA Test

Section	Number of items			
	Grade K	Grade 1	Grade 2	Grade 3
Counting and Cardinality	1	0	0	0
Word Problems	3	4	5	5
Number Relations	4	6	7	8
Fractions	2	4	4	5
Basic Number Facts	4	4	4	2
Computation	3	4	5	8
Total	17	22	25	28

To support vertical scaling across grade levels, at least two items (called anchor items) within each section on test forms at adjacent grade levels were identical by design. This goal was achieved in all cases for the items in the counting and cardinality section, which was presented only to grade K. In general, the items were intended to be sequenced from easier to more difficult within each subsection.

1.1.1. Counting and Cardinality

Table 1.2 provides an overview of the Counting and Cardinality items by grade level. Counting and Cardinality items were not presented to grade 1 through 3 students to ensure the appropriate level of difficulty for each grade level.

Table 1.2. Items in the Counting and Cardinality Section

Bank item number	Response format	Grade K	Grade 1	Grade 2	Grade 3
CT1.4.S1	SR				

Note. SR = selected response.

1.1.2. Word Problems

Table 1.3 shows the types of problems included in the Word Problems section of the Spring 2019 K–3 EMSA test at each grade level. The three-letter abbreviations for the problem types correspond to the names of word problems as defined by Carpenter, Fennema, Franke, Levi, & Empson (2015). The numbers correspond to the two numbers given in the problem. The anchor set for the grades K–1 includes two items, that for the grades 1–2 includes two, and for grades 2–3 includes four.

Table 1.3. Items in the Word Problems Section

Bank item number	Response format	Grade K	Grade 1	Grade 2	Grade 3
WP22.S1	SR				
WP34.S1	SR				
WP24.C	CR				
WP4.2.C	CR				
WP26.C	CR				
WP16.C	CR				
WP35.C	CR				
WP43.C	CR				
WP55.C	CR				

Note. CR = constructed response; SR = selected response. For a full list of the problem-type abbreviations, see the List of Abbreviations or Carpenter et al. (2015).

The two *multiplication-grouping* (MG) problems and the *measurement-division* (MD) problem might be considered to be beyond the scope of the content of the CCSS-M at grades K, 1, and 2. We include them in the tests because abundant empirical evidence demonstrates that students at these grade levels can solve these types of problems (Carpenter, Ansell, Franke, Fennema, & Weisbeck, 1993; Turner & Celedón-Pattichis, 2011; Verschaffel, Greer, & De Corte, 2007). Moreover, the focus on place value and the base-ten structure of the number system in the mathematics curriculum standards at the early elementary level involves grouping situations—with a particular focus on groups of ten—consistent with MG and MD problems (Carpenter et al., 2015).

All the items in this section at each grade level were read aloud twice to the students as per the administration instructions (see Appendices E, F, G, and H) to help focus the assessment on students' mathematics achievement and lessen the potential effect of their respective reading or listening comprehension ability.

1.1.3. Number Relations Section

The Number Relations section contained items presented to measure students' understanding of the equals sign. The anchor set for grades K–1 included four items, that for grades 1–2 included six, and that for grades 2–3 included seven. Four items in this section ( ) were included at all four grade levels, thereby creating a potential set of anchor items among all four grade levels. The first two items were included as warm-up items to assist students with understanding the intent of the items in this section. All selected-response items in this section were presented together with no more than two items per page. The constructed-response items were embedded within the computation and basic number facts section.

Table 1.4. Items in the Number Relations Section

Bank item number	Response format	Grade K	Grade 1	Grade 2	Grade 3
ES5.S1	SR				
ES8.S1	SR				
ES15.S1	SR				
ES1.3.S1	SR				
ES18.S1	SR				
ES3.3.C	CR				
ES.27.C	CR				
ES21.C	CR				

Note. CR = constructed response; SR = selected-response.

The selected-response items in this section are read aloud to students. For the first five items shown in Table 1.4 students are asked to circle “yes” if the equation was correct or to circle “no” if the equation was not correct.

1.1.4. Fractions Section

Table 1.5 provides an overview of the items in the Fractions section of the Spring 2019 K–3 EMSA test at each grade level.

Table 1.5. Items in the Fractions Section

Bank item number	Response format	Grade K	Grade 1	Grade 2	Grade 3
FPW2.2.S1	SR				
FPW3.3.S1	SR				
FPW3.5.S1	SR				
FPW14.S1	SR				
FNL2.1.S1	SR				
FWP7.S1	SR				

Note. CR = constructed response; SR = selected-response.

This section of the assessment was intended to gather information about (a) students' understanding of fractions as they are conceptualized as a relationship between parts and wholes and (b) students' understanding of the number line as a representational tool for rational numbers.

The items in this section involve concepts such as (a) fractions are always defined with respect to a referent unit, the whole; (b) fractions consist of a certain number of equal or same-size pieces of a

whole or unit; (c) vulgar fractions can be built from the iteration of unit fractions; and (d) the inverse relation between the magnitudes of the denominator and the corresponding fractional part.

The majority of the items in this section were presented in a selected-response format. Three used a constructed-response format, and three required students to draw a circle around shapes that represented a certain fraction. As indicated in Table 1.5, the anchor set for grades K–1 included two items, that for grades 1–2 included four, and that for grades 2–3 included two.

1.1.5. Basic Number Facts Section

Table 1.6 provides an overview of the items in the Basic Number Facts section of the Spring 2019 K–3 EMSA test at each grade level. In this section, the anchor sets for grades K–1 included two items, that for grades 1–2 included four, and that for grades 2–3 included two. One item was included across all grade levels.

Table 1.6. Items in the Basic Number Facts Section

Bank item number	Response format	Grade K	Grade 1	Grade 2	Grade 3
CBF15.S2	SR				
CBF21.S2	SR				
CBF3.S2	SR				
CBF10.S2	SR				
CBF35.C	CR				
CBF5.C	CR				

Note. CR = constructed response; SR = selected-response.

1.1.6. Computation Section

The final section of the test was designed to measure students' ability to compute sums and differences with basic facts and higher numbers. The problems in this section were more varied, as grade-2 students are typically more proficient with computing sums and differences of greater numbers than kindergarten or grade-1 students. Grade-3 student test forms included operations involving multiplication of whole numbers and addition and subtraction of fractional numbers.

With the exception of grade K, students were given directions for completing some of the items on their own. They were told that they would work on some addition and subtraction problems on their own for grades K, 1, and 2, and addition, subtraction, and multiplication problems for grade 3. Students were informed they would solve them at their own pace. According to the script in the administration guide, the students were encouraged to look closely at the symbol to decide whether each problem involved addition, subtraction, or multiplication.

Table 1.7 provides an overview of the items in the Computation section of the Spring 2019 K–3 EMSA test at each grade level. In this section, the anchor set for grades K–1 included three items, that for grades 1–2 included three items, that for grades 2–3 included five items. Because of the very wide range of ability in computation between grade K and grade 2 students, no items were identical across all three grade levels in this section.

Table 1.7. Items in the Computation Section

Bank item number	Response format	Grade K	Grade 1	Grade 2	Grade 3
CWN1.1.S2	SR				
CWN2.1.S2	SR				
CWN4.1.C	CR				
CWN3.4.C	CR				
CWN25.C	CR				
CTWN6.1.C	CR				
CWN42.C	CR				
CWN33.2.C	CR				
CWN49.C	CR				
CF7.C	CR				
CWN50.C	CR				

Note. CR = constructed response; SR = selected-response. Items in bold-face font were read aloud to students.

1.1.7. Detailed Test Blueprint

Table 1.8 provides a detailed blueprint showing the items in each of the six sections of the test. Items displayed with a strikethrough were on the test form but were removed from the final scale as a result of poor item statistics. See sections 3 and 4 of the present report for more information on the review and analysis of the individual items.

Table 1.8. Detailed Test Blueprint for the Spring 2019 K–3 EMSA

Item description	Bank item number	Test item number			
		Grade K	Grade 1	Grade 2	Grade 3
Counting and Cardinality					
Determine how many apples (12)	CT1.4.S1	1			
Word Problems					
	WP22.S1	2			
	WP34.S2	3	1		
	WP24.C	4	2		
	WP4.2.C		3	1	
	WP26.C			2	1
	WP16.C			3	2
	WP35.C				3
	WP43.C		4	4	4
	WP55.C			5	5
Number Relations					
	ES5.S1	5	5	6	6
	ES8.S1	6	6	7	7
	ES15.S1	7	7	8	8
	ES1.3.S1	8	8	9	9
	ES18.S1		9	10	10
	ES3.3.C		20	21	20
	ES.27.C				24
	ES21.C			24	26
Fractions					
	FPW2.2.S1	9	10	11	
	FPW3.3.S1	10	11	12	
	FPW3.5.S1		12	13	11
	FPW14.S1		13	14	12
	FNL2.1.S1				13
	FWP7.S1				14

(Continued)

Item description	Bank item number	Test item number			
		Grade K	Grade 1	Grade 2	Grade 3
	FPW7.C				15
	<i>Basic Number Facts</i>				
	CBF15.S2	11			
	CBF21.S2	12			
	CBF3.S2	13	14	15	
	CBF10.S2		~~15~~	~~16~~	~~16~~
	CBF35.C	16	18	~~19~~	~~19~~
	CBF5.C		~~19~~	~~20~~	
	<i>Computation</i>				
	CWN1.1.S2	14			
	CWN2.1.S2	15	16		
	CWN4.1.C		17	17	17
	CWN3.4.C			18	18
	CWN25.C	17	21		
	CTWN6.1.C		22	22	21
	CWN42.C			23	22
	CWN33.2.C			25	27
	CWN49.C				23
	CF7.C				25
	CWN50.C				28
Items on test form		17	22	25	28
Items in final scale		15	18	20	24

Note. The items with strikethrough font are the items that were on the test form but were removed from the final scale.

1.2. Description of the Sample and Setting

Students in the sample attended schools where their teachers had volunteered to participate in a research study examining the effects of teacher and student mathematics anxiety on students' mathematical performance. Tests forms were delivered to schools by project staff in mid-April 2019. In the field tests reported here, the students' classroom teachers administered the tests during the testing window of April 15–May 3.

The analytic sample included 2,538 students drawn from 238 classrooms in 29 schools representing 3 Florida public school districts. Table 1.9 provides descriptive statistics for the student sample from data available to us at the time of this report.

Table 1.9. Demographic Characteristics of the Students in the Spring 2019 Field Test of the K–3 EMSA Tests

Student characteristic	Number				Overall sample (n = 2538)
	Grade K (n = 568)	Grade 1 (n = 636)	Grade 2 (n = 650)	Grade 3 (n = 684)	
Gender					
Male	267	294	310	315	1186
Female	286	284	306	333	1209
Unknown	15	58	34	36	143
Race					
White	302	313	322	325	1262
Black	115	120	128	172	535
Asian	23	30	28	21	102
Other	19	3	12	11	45
Unknown	109	170	160	155	594
Ethnicity					
Hispanic	90	110	130	112	442
Non-Hispanic	457	469	477	535	1938
Unknown	21	57	43	37	158

Note. ELL = English language learner. SWD = Students with disabilities. Individual student demographic characteristics, such as exceptionality, or eligibility for free or reduced-price lunch, were not available for many students at the time of the present report was written. Some of the proportions do not sum to 1.00 because of rounding errors.

In the 2018–19 school year, the Mathematics Florida Standards defined the official set of standards for mathematics in grades K–12 (Florida Department of Education, 2014). The CCSS-M (NGACBP & CCSSO, 2010) and the Mathematics Florida Standards are similar to one another but are not identical at these grade levels. No statewide assessment of student mathematics achievement in grades K–2 is conducted in Florida, but some individual districts use district-selected assessment tools to monitor progress of K–2 students. The Florida State Board of Education requires public schools to administer the Florida Standards Assessment (FSA) to students in grade 3 for the purpose of school accountability.

2. Test Development, Scoring, and Data-Entry Procedures

2.1. Content

The content standards at grades K, 1, 2, and 3 in the CCSS-M (NGACBP & CCSSO, 2010) and Mathematics Florida Standards (Florida Department of Education, 2014) provide guidelines for content specifications. Overall, the focus of the test is on number (whole and fractions), operations, and equality. It includes several items designed to favor students who have a solid grasp of place-value concepts. The highest number on the grade K test is 20, that on the grade 1 test is 37, that on the grade 2 test is 686, that on the grade 3 test is 10,000. Computation items presented symbolically involve applying the addition, subtraction, or multiplication operation between exactly two positive integers. Problems involving subtraction result in a difference with a positive, integer value. Word problems involve additive situations as well as grouping situations that could be solved by multiplication, division, addition, counting strategies, or direct place-value understanding (Carpenter, Fennema, Franke, Levi, & Empson, 1999; Carpenter et al., 2015).

Test design involved finding an optimal point at the intersection of two potentially competing goals: (1) to sample a range of difficulty of problems and cognitive demand to reflect students' mathematical understanding and the learning goals outlined in grades K, 1, 2, and 3 in the CCSS-M and the Mathematics Florida Standards and (2) to minimize the testing burden on teachers and students.

2.2. Instrument Development Process

The development process for the Spring 2019 K–3 EMSA tests consisted of the following phases:

1. Review of the content and psychometric properties of the 2014 and 2015 MPAC Interviews (Schoen et al., 2016b; 2016c); the Fall 2013, Fall 2014, Fall 2015, Fall 2018, and Spring 2016 EMSA test items (Schoen et al., 2016a, 2016b; Schoen et al., 2017; 2018); and the Early Fractions Test (Schoen et al., 2017).
2. Review of content expectations for grades K, 1, 2, and 3 in the Mathematics Florida Standards (Florida Department of Education, 2014) and CCSS-M (NGACBP & CCSSO, 2010) and alignment with available items in the EMSA item bank
3. Development of the first written draft of the test blueprint
4. Internal review of the draft test blueprint by members of the evaluation team and external review by experts in mathematics and mathematics education
5. Revision of the test blueprint based on feedback
6. Development of the first written draft of the test form for grades K, 1, 2, and 3 and corresponding scoring procedures
7. Analysis of the frequency of correct response position and distribution of correct response positions across each grade-level test
8. Review of the draft test forms, including formatting, editing, and proofing
9. Development of administration instructions
10. Proofreading of test and administration instruction forms
11. Printing and distribution of test booklets, administration instructions, and ancillary documents.

Test items from several tests previously administered in the fall or spring with grade K, 1, 2, and 3 students informed the test in several ways. Items with poor psychometric statistics from previous field tests were not used. Many of the items on the previous field tests had an open-ended, constructed-response format. The students' responses in the open-ended format on previous test forms and pilot studies informed the determination of the set of response options in the selected-response items of the Spring 2019 K–3 EMSA tests. In general, the four most frequently provided responses were used as the four response options on selected-response items. When we did not have data to support the conclusion that four responses would represent at least 80% of the responses that students would present in a constructed-response format, we presented the item as a constructed-response item on the Spring 2019 EMSA.

During the process of expert review, test items were reviewed for content accuracy as well as potential bias and sensitivity in an effort to neutralize effects of nonmathematical vocabulary terms, grammatical structure, or reading levels. Whenever possible, word problems were written to avoid the use of keywords (e.g., altogether, in all, left).

2.3. Test Design and Assembly

A sample item with an example of responses is provided on the first page of the test for the administrator to use in demonstrating how students are expected to respond (e.g., by completely filling the oval). The students were directed to fill in the ovals (which were called “bubbles” in the administration guide) below their answer choices. Ovals were centered beneath the corresponding response options, and responses were centered horizontally across the page. The response options for selected-response items with numerical solutions were ordered left-to-right from least to greatest. Copies of the grades K, 1, 2, and 3 tests are presented in Appendices A, B, C, and D respectively. Table 2.1 shows how many selected-response and constructed-response items were included in each category for each grade level.

Table 2.1. Number of Constructed- and Selected-Responses on Spring 2019 K-3 EMSA

Item format	Number of items			
	Grade K	Grade 1	Grade 2	Grade 3
<i>Counting and Cardinality</i>				
Constructed-response				
Selected-response	1			
Four response options	1			
Two response options				
<i>Word Problems</i>				
Constructed-response	1	3	5	5
Selected-response	2	1		
Four response options	2	1		
<i>Number Relations</i>				
Constructed-response		1	2	3
Selected-response	4	5	5	5
Four response options				
Three response options				
Two response options	4	5	5	5
Circle shapes				
<i>Fractions</i>				
Constructed-response				1
Selected-response	2	4	4	4
Four response options				2
Three response options		1	1	1
Two response options				
Circle shapes	2	3	3	1
<i>Basic Number Facts</i>				
Constructed-response	1	2	2	1
Selected-response	3	2	2	1
Four response options	3	2	2	1
<i>Computation</i>				
Constructed-response	1	3	5	8
Selected-response	2	1		
Four response options	2	1		

No pictures or images appeared on the page apart from the page-numbering system, the text of the problem, the numerals comprising the response options, the four ovals that students use to indicate their responses, and the boxes in which students recorded their responses on constructed-response

items. Plenty of empty space was available on the page for students to draw or record their thoughts as necessary.

Most of the items in the Computation and Basic Number Facts sections and three of the items in the Number Relations section consist of items presented as open equations. Students were instructed to write their answers in the empty box representing the unknown value in the equation to be solved.

In the Word Problems section, only one problem is displayed per page so that students will not record their answers in the wrong places or be overwhelmed by too much visual information. For grades 1, 2, and 3, in the Computation, Basic Number Facts, and Number Relations sections, several items are presented per page over multiple pages. The grade K test contained no more than two items per page. The grammar used in word problems was reviewed by people with expertise in teaching emergent bilingual students. Arial font in large type size was used throughout.

For grades K, 1, and 2 pages were identified by a series of child-friendly images rather than page numbers. Figure 1 provides one example of these images. The large and easily distinguished image is also useful for the test administrator to use as a way to verify from across the room that students have turned to the correct page. Page numbers in large font were used in upper corners of the grade 3 test.

Figure 2.1. One of the images used in place of page numbers.

During the test development, careful consideration was given to the distribution of correct-response positions across each test form. Table 2.1 provides the number of times the correct answer was in each position at each grade level.

Table 2.2. Number of Times the Correct Answer Appears in Each Position

Grade level	A	B	C	D	Yes	No
K	1	3	2	2	3	1
1	0	4	0	1	4	1
2	0	3	0	0	4	1
3	0	2	1	1	4	1

2.4. Test Production

The tests, administration guides, and consent forms were printed on blue, 20-pound paper at Florida State University and distributed to the participating schools. All test forms were printed double-sided. Blue paper was used primarily to prevent bleed through from front-to-back and for a secondary purpose of having a calming effect on the examinees.

Test administration guides were created for each test and were specific to each grade-level. The administration guide was repeatedly reviewed, edited, and proofread by research project staff during the test-development process. Drafts of administration guides were printed, and data-entry personnel proofed each grade-level administration guide by administering the assessment to other data-entry personnel to ensure that the prompts were correctly written in the script. Data-entry personnel

provided feedback to the authors, and revisions were made before the final version was printed. The research study associated with this administration asked teachers to administer two other assessments, and these directions were included in the administration guides. To prevent confusion on the part of test administrators, the scripts for each assessment were printed on the same colored paper as the student assessment (i.e., blue for the EMSA, green for the study survey, etc.).

2.5. Administration Protocol for the Spring 2019 EMSA K–3 Test

Each participating teacher was provided with a test packet containing

- Test-administration guide (for the corresponding grade level)
- Class set of student tests
- Student information sheet
- Parental consent forms (for those students who had not already submitted consent)
- Student assent forms (for those students who had not already given assent)

These materials were distributed to the main offices at school sites during the mid-April. They were distributed to the participating teachers through the main office personnel or principal-appointed designee. Teachers were instructed to administer the tests between April 15 and May 3.

Test administrators were given detailed instructions (which appear in Appendices E, F, G, and H) on how to administer the test, including a script to use during administration. Test administrators were asked to write students' names on the front covers of the tests that did not have a label with the students' name already printed to increase legibility and accuracy in data entry. They were also instructed to permit students to use manipulable materials if that was common practice in their classrooms. The Spring 2019 K–3 EMSA test administration guides provided an overview of the tests, described the administration process and directions, explained how to submit completed tests, and provided a full script to be read verbatim during administration of the test.

The K–3 EMSA tests were designed to be administered in a whole-group setting in a paper-pencil format. Test administrators were given a guide explaining how to administer the tests, along with a script to use while doing so. Depending on the type of item, students were instructed to record their responses by shading bubbles to indicate their responses to selected-response items, write their responses to constructed-response items in boxes provided on the paper forms, or draw rings around images to indicate their responses.

The test administrators were instructed to read each problem in the Counting and Cardinality, Word Problems, and Fractions sections aloud to students at all grade levels to ensure that students' reading comprehension did not interfere with mathematics thinking. Some problems in Basic Number Facts, Computation, and Number Relations sections were read aloud twice to students. The test administrator was instructed to read all items on the grade K test aloud to students. In grades 1 through 3, the test administrator was instructed to read directions for the Computation section and to invite students to read the individual items on their own and work at their own pace. The test administrators were told that they could reread a problem on request, but they were instructed always to read the entire problem exactly as written and to refrain from reading just a portion of it. Almost all items in the Computation section and three items in the Number Relations section used a constructed-response format. With the exception of grade K, students completed these items independently and were instructed to write the number in the box that would make the equation correct.

During test administration, students record their responses with pencils directly on the test booklet. Students were allowed to use blank space provided in the test booklet if they needed space to write. In most cases, the students' classroom teacher administered the test.

The Spring 2019 K–3 EMSA test was not a timed test. Test administrators were informed that the test required approximately 45 minutes to administer, but administration time was allowed to vary across settings. Test administrators were encouraged to provide students with sufficient time to complete each item on the test to their own satisfaction.

The testing conditions for the Spring 2019 K–3 EMSA were expected to be held consistent with the testing conditions used in other student assessments administered in the teacher's classroom. For example, students should separate their desks or use student "privacy folders" if that is what they usually do. In addition, students were permitted to use mathematics manipulatives or other tools during the Spring 2019 K–3 EMSA if they were ordinarily permitted to do so in that particular classroom.

If students required testing accommodations resulting from IEP, ELL, or 504 plans, then the teacher was expected to provide any and all required accommodations for those individual students and to document the accommodation on the student information sheet. The test was not designed to be timed, so test administrators were instructed to allow students adequate time to answer all of the questions.

Student participation in the study was overseen by the human subjects committee (i.e., institutional review board) at Florida State University and by the school districts of participating students. Parental consent forms were distributed and collected in fall 2018. The test packets sent to classrooms contained consent letters and forms to be distributed to those students who had not returned consent forms at the time of the Fall 2018 EMSA test administration.

Participating teachers received a class roster in the test packet. The roster indicated consent status of students in the class who were known to be permitted to participate in the study. Upon conclusion of test administration, teachers were instructed to submit all testing materials (test administration guide, used and unused student test booklets, student information sheet, and parental consent forms when applicable) to their principals or designees. Teachers were asked to return only test booklets completed by those students with parental consent. The principal or designee placed the testing materials in the main office at the front desk for pickup. Members of the project team picked up test materials during the third week of May 2019.

Test administrators were instructed to insert completed tests into an opaque envelope, seal it, and deliver it to the front office for research-project personnel to pick up during a window of time outlined in the administration instructions.

Teachers who presented extenuating circumstances to the research team and did not administer the test during the administration window or missed the materials pickup date were handled on a case-by-case basis with respect to when to administer the test and arrangement for materials pickup. Two teachers were granted a time extension for materials pickup. The date of test administration was not used as a factor in data modeling.

Teacher administration of the test presents the potential for breaches in test security. These were not high-stakes tests, so strict security was not a high priority. In this case, teachers and schools were trusted to administer the tests in accordance with the instructions.

2.6. Data Entry and Verification Procedures

Research assistants transferred student responses on the paper-based test forms into forms hosted on a FileMaker Pro database (FileMaker Pro, Version 16). Data for constructed-response items were manually keyed into the FileMaker forms exactly as written. As a strategy to minimize data-entry errors through typographical errors, response fields for selected-response items were validated to allow only the codes for offered responses, as well as codes for missing or uninterpretable responses.

Sometimes, students neglected to fill in the bubble corresponding to their responses. In these instances, research assistants examined written work provided by the student to determine whether a clear intended response seemed to have been given. When a clearly intended response matched one of the response options given, it was then coded as though the student had filled in the bubble. For example, when an examinee wrote a numeral in the position of the unknown variable in line with the equation and did not fill the bubble below the corresponding numeral, the response was entered as though the bubble had been filled. When the intended response did not match one of the response options presented, when multiple responses were selected, or when no intended response was evident, the response was coded as unclear intent (UI). If no attempted response was evident, the response was coded as no response (NR).

Research assistants were given the task of interpreting both the student's handwriting and the student's intent, with the goal of entering the student's intended response exactly as it was written. Most of the students' responses to the constructed-response items were entered as the student wrote them, but research assistants could choose one of the missing-item codes, UI and NR, when applicable.

To verify the accuracy of the entered data, research assistants entered the results from 11% of the tests twice. These tests were selected through a random sample of participating classes, stratified by grade level. When possible, two different assistants entered data for the first and second entries. This procedure ensured that the double-entry comparison was performed on a sample representing all five grade levels and all the data-entry personnel. These separate records were compared and found to have an overall 99.8.0% agreement on items entered.

2.7. Item-scoring Procedures

Responses were scored according to the scoring guide provided in Appendix I of this report. Although each item on the Spring 2019 EMSA was designed to have a unique, correct solution, examinees could generate unexpected responses that could be judged to be correct, such as mathematically equivalent expressions (e.g., 5, $4 + 1$). Student responses to each constructed-response item were aggregated and reviewed by an adjudication committee composed of research assistants and experts in mathematics to identify any unanticipated—but mathematically correct—responses. Item responses coded as NR, UI, and MR were scored as incorrect. All available data were used in the subsequent analyses.

3. Data and Scoring

3.1. Missing Response Data

EMSA Spring 2019 test kindergarten, grade-1, grade-2, and grade-3 results are presented here, including: 573 students and 17 items in kindergarten, 638 students and 22 items in grade 1, 655 students and 25 items in grade 2, and 690 students and 28 items in grade 3.

Initially, item responses were coded as correct, incorrect, or missing. The items and scoring key are provided in Appendix I. After missing data were dealt with, item responses were coded as correct or incorrect.

We determined that an examinee must provide at least some type of response to at least 70% of the items for a test to be valid. Other examinees were excluded from further analysis. For examinees retained, missing responses were coded as incorrect. These operations were carried out before any item was dropped.

Five students were excluded from kindergarten, 2 from grade 1, 5 students from grade 2, and 6 from grade 3. After the exclusions, 568 students remained in kindergarten, 636 in grade 1, 650 in grade 2, and 684 in grade 3. Tables 3.1–3.4 show the frequency and percentage of missing responses for each grade level.

Table 3.1. Missing Response Frequency in Kindergarten Sample

No. of missing responses	Frequency	%	Cumulative %
0	501	87.4	87.4
1	44	7.7	95.1
2	21	3.7	98.8
3	2	0.3	99.1
6	3*	0.5	99.7
7	1*	0.2	99.8
8	1*	0.2	100
Total	573	100	

Note. No. of missing responses = the number of missing responses for a given student in the sample; frequency = the number of students with that number of missing responses; % = the percentage of students with that number of missing responses; cumulative % = cumulative percentage of students with that number of missing responses. The total number of items in kindergarten was 17.

*Students excluded from the analysis.

Table 3.2. Missing Response Frequency in Grade 1 Sample

No. of missing responses	Frequency	%	Cumulative %
0	580	90.9	90.9
1	33	5.2	96.1
2	15	2.4	98.4
3	3	0.5	98.9
4	3	0.5	99.4
6	2	0.3	99.7
7	2*	0.3	100
Total	638	100	

Note. Column headings as in Table 3.1. The total number of items in grade 1 was 22.

*Students excluded from the analysis.

Table 3.3. Missing Response Frequency in Grade 2 Sample

No. of missing responses	Frequency	%	Cumulative %
0	600	91.6	91.6
1	27	4.1	95.7
2	10	1.5	97.3
3	6	0.9	98.2
4	6	0.9	99.1
7	1	0.2	99.2
10	1*	0.2	99.4
14	1*	0.2	99.5
17	1*	0.2	99.7
25	2*	0.3	100
Total	655	100	

Note. Column headings as in Table 3.1. The total number of items in grade 2 was 25.

*Students excluded from the analysis.

Table 3.4. Missing Response Frequency in Grade 3 Sample

No. of missing responses	Frequency	%	Cumulative %
0	612	88.7	88.7
1	39	5.7	94.3
2	12	1.7	96.1
3	7	1	97.1
4	5	0.7	97.8
5	6	0.9	98.7
6	2	0.3	99
8	1	0.1	99.1
12	1*	0.1	99.3
14	1*	0.1	99.4
15	2*	0.3	99.7
22	1*	0.1	99.9
25	1*	0.1	100
Total	690	100	

Note. Column headings as in Table 3.1. The total number of items in grade 3 was 28.

*Students excluded from the analysis.

3.2. Initial Item Analysis and Item Removal

Two warm-up items—very easy items that were intended to introduce examinees to the format of the items involving true/false questions about the equals sign but not intended to be used for scoring purposes—were removed from the tests for further analysis. As a result, items 5 and 6 in the kindergarten and grade 1 tests and items 6 and 7 in the grade 2 and 3 tests were removed.

After the warm-up questions were removed, we conducted analyses based on classical testing theory (CTT) and item response theory (IRT) as well as content analysis. We decided to drop items 15 and 19 from grade 1; items 16, 19, and 20 from grade 2; and items 16 and 19 from grade 3, because the content analysis revealed oversampling of items measuring the same content domain. We dropped those items in the same content domain which had close item difficulty values according to the CTT and IRT item parameter estimates.

In the final scale, items 5 and 6 were excluded from kindergarten; items 5, 6, 15, and 19 from grade 1; items 6, 7, 16, 19, and 20 from grade 2; and items 6, 7, 16, and 19 from grade 3, leaving 15 items in kindergarten, 18 in grade 1, 20 in grade 2, and 24 in grade 3. Tables 3.5–3.8 show item indexing and scoring of test-form and final-scale items for each grade level.

Table 3.5. Item Indexing and Scoring for Both Test-Form and Final-Scale Formats in Kindergarten

Bank item number	Test-form item #	Final-scale item #
CT1.4.S1	1	1
WP22.S1	2	2
WP34.S2	3	3
WP24.C	4	4
ES5.S1	5	
ES8.S1	6	
ES15.S1	7	5
ES1.3.S1	8	6
FPW2.2.S1	9	7
FPW3.3.S1	10	8
CBF15.S2	11	9
CBF21.S2	12	10
CBF3.S2	13	11
CWN1.1.S2	14	12
CWN2.1.S2	15	13
CBF35.C	16	14
CWN25.C	17	15

Note. Test-form Item # = the item index from the original test; Final-scale item # = the newly generated item number after items 5 and 6 were excluded. All items were scored dichotomously: 0, 1.

Table 3.6. Item Indexing and Scoring for Both Test-Form and Final-Scale Formats in Grade 1

Bank item number	Test-form item #	Final-scale item #
WP34.S2	1	1
WP24.C	2	2
WP4.2.C	3	3
WP43.C	4	4
ES5.S1	5	
ES8.S1	6	
ES15.S1	7	5
ES1.3.S1	8	6
ES18.S1	9	7
FPW2.2.S1	10	8
FPW3.3.S1	11	9
FPW3.5.S1	12	10
FPW14.S1	13	11
CBF3.S2	14	12
CBF10.S2	15	
CWN2.1.S2	16	13
CWN4.1.C	17	14
CBF35.C	18	15
CBF5.C	19	
ES3.3.C	20	16
CWN25.C	21	17
CTWN6.1.C	22	18

Note. Test-form Item # = the item index from the original test; Final-scale item # = the newly generated item number after items 5, 6, 15, and 19 were excluded. All items were scored dichotomously: 0, 1.

Table 3.7. Item Indexing and Scoring for Both Test-Form and Final-Scale Formats in Grade 2

Bank item number	Test-form item #	Final-scale item #
WP4.2.C	1	1
WP26.C	2	2
WP16.C	3	3
WP43.C	4	4
WP55.C	5	5
ES5.S1	6	
ES8.S1	7	
ES15.S1	8	6
ES1.3.S1	9	7
ES18.S1	10	8
FPW2.2.S1	11	9
FPW3.3.S1	12	10
FPW3.5.S1	13	11
FPW14.S1	14	12
CBF3.S2	15	13
CBF10.S2	16	
CWN4.1.C	17	14
CWN3.4.C	18	15
CBF35.C	19	
CBF5.C	20	
ES3.3.C	21	16
CTWN6.1.C	22	17
CWN42.C	23	18
ES21.C	24	19
CWN33.2.C	25	20

Note. Test-form Item # = the item index from the original test; Final-scale item # = the newly generated item number after items 6, 7, 16, 19, and 20 were excluded. All items were scored dichotomously: 0, 1.

Table 3.8. Item Indexing and Scoring for both Test-Form and Final-Scale Formats in Grade 3

Bank item number	Test-form item #	Final-scale item #
WP26.C	1	1
WP16.C	2	2
WP35.C	3	3
WP43.C	4	4
WP55.C	5	5
ES5.S1	6	
ES8.S1	7	
ES15.S1	8	6
ES1.3.S1	9	7
ES18.S1	10	8
FPW3.5.S1	11	9
FPW14.S1	12	10
FNL2.1.S1	13	11
FWP7.S1	14	12
FPW7.C	15	13
CBF10.S2	16	
CWN4.1.C	17	14
CWN3.4.C	18	15
CBF35.C	19	
ES3.3.C	20	16
CTWN6.1.C	21	17
CWN42.C	22	18
CWN49.C	23	19
ES.27.C	24	20
CF7.C	25	21
ES21.C	26	22
CWN33.2.C	27	23
CWN50.C	28	24

Note. Test-form Item # = the item index from the original test; Final-scale item # = the newly generated item number after items 6, 7, 16, and 19 were excluded. All items were scored dichotomously: 0, 1.

4. Dimensionality Analysis

4.1. Parallel Analysis

Parallel Analysis (PA) first proposed by Horn (1965), is a method used to determine the number of factors to retain, based on either principal-components or principal-factor analysis. It calls for comparison of the eigenvalues of the original data to the mean of eigenvalues generated from random correlation matrices that have properties similar to those of the original data. Zwick and Velicer (1986) compared Horn's PA with rule-of-thumb procedures and found that the PA provided more accurate results, but studies showed that Horn's version of PA had tendencies to retain minor/nuisance factors (Zwick & Velicer, 1986; Glorfeld, 1995). Glorfeld (1995) proposed to use the percentiles of the distributions of the eigenvalues generated from a large number of random correlation matrices. The results showed that this modified version of the PA avoided overextraction of the number of factors, which Zwick and Velicer (1986) had pointed out.

We conducted the PA to determine the number of factors of the test at each grade level. The *psych* (Revelle, 2020) program in R 4.0.2 (R Core Team, 2020) was used to perform the analysis. Figures 4.1–4.4 show the results of the PA of the test at each grade level. In the figures, the horizontal axis represents the number of components, which also indicates the number of variables, whereas the vertical axis shows the eigenvalues of the principal components. The dotted line shows the number of factors that PA suggested being retained. The results suggested a single, dominant factor for the test at each grade level.

Figure 4.1. Parallel analysis scree plot of kindergarten test.

Figure 4.2. Parallel analysis scree plot of grade-1 test.

Figure 4.3. Parallel analysis scree plot of grade-2 test.

Figure 4.4. Parallel analysis scree plot of grade-3 test.

4.2. Confirmatory Factor Analysis

Confirmatory factor analysis reveals whether the relationships among the test items (i.e., covariance matrix) are a function of a hypothesized factor structure. In addition to the PA, we performed the confirmatory factor analysis using Mplus 8.1 (Muthén & Muthén, 2017). Weighted least squares mean and variance adjusted (Muthén, du Toit, & Spisic, 1997) estimation based on tetrachoric correlation was used. Fit indices of comparative fit index, root mean square error of approximation (RMSEA), and standardized root mean residual along with standardized item loadings from each grade-level test result are presented in Tables 4.1–4.5.

Table 4.1. Model-fit Statistics of the Test in Each Grade Level

Fit index	Kindergarten	Grade 1	Grade 2	Grade 3
CFI	0.981	0.957	0.950	0.954
RMSEA	0.029	0.037	0.053	0.047
SRMR	0.064	0.074	0.085	0.077

Note. CFI = comparative fit index; RMSEA = root mean square error of approximation; SRMR = standardized root mean residual.

Table 4.2. Factor Loadings from Kindergarten Test confirmatory factor analysis (CFA)

Bank item number	Final-scale item #	Loading	s.e.
CT1.4.S1	1	0.39	0.09
WP22.S1	2	0.58	0.06
WP34.S2	3	0.45	0.05
WP24.C	4	0.73	0.04
ES15.S1	5	0.53	0.05
ES1.3.S1	6	0.23	0.06
FPW2.2.S1	7	0.56	0.05
FPW3.3.S1	8	0.43	0.06
CBF15.S2	9	0.63	0.06
CBF21.S2	10	0.82	0.04
CBF3.S2	11	0.77	0.04
CWN1.1.S2	12	0.61	0.06
CWN2.1.S2	13	0.65	0.04
CBF35.C	14	0.66	0.04
CWN25.C	15	0.71	0.04

Note. Final-scale item # = the newly generated item number after excluding items 5 and 6; Loading is the standardized value of the item loadings.

Table 4.3. Factor Loadings from Grade-1 Test CFA

Bank item number	Final-scale item #	Loading	s.e.
WP34.S2	1	0.71	0.04
WP24.C	2	0.79	0.03
WP4.2.C	3	0.67	0.04
WP43.C	4	0.73	0.04
ES15.S1	5	0.52	0.05
ES1.3.S1	6	0.51	0.05
ES18.S1	7	0.49	0.05
FPW2.2.S1	8	0.71	0.04
FPW3.3.S1	9	0.47	0.05
FPW3.5.S1	10	0.51	0.05
FPW14.S1	11	0.30	0.06
CBF3.S2	12	0.48	0.07
CWN2.1.S2	13	0.60	0.05
CWN4.1.C	14	0.72	0.04
CBF35.C	15	0.46	0.06
ES3.3.C	16	0.61	0.05
CWN25.C	17	0.58	0.05
CTWN6.1.C	18	0.42	0.05

Note. Final-scale item # = the newly generated item number after excluding items 5, 6, 15, and 19; Loading is the standardized value of the item loadings.

Table 4.4. Factor Loadings from Grade-2 Test CFA

Bank item number	Final-scale item #	Loading	s.e.
WP4.2.C	1	0.65	0.04
WP26.C	2	0.58	0.04
WP16.C	3	0.65	0.04
WP43.C	4	0.64	0.04
WP55.C	5	0.54	0.05
ES15.S1	6	0.58	0.06
ES1.3.S1	7	0.58	0.05
ES18.S1	8	0.77	0.03
FPW2.2.S1	9	0.55	0.06
FPW3.3.S1	10	0.40	0.05
FPW3.5.S1	11	0.35	0.05
FPW14	12	0.27	0.05
CBF3.S2	13	0.58	0.08
CWN4.1.C	14	0.77	0.05
CWN3.4.C	15	0.65	0.04
ES3.3.C	16	0.90	0.02
CTWN6.1.C	17	0.50	0.04
CWN42.C	18	0.63	0.04
ES21.C	19	0.99	0.02
CWN33.2.C	20	0.44	0.05

Note. Final-scale item # = the newly generated item number after excluding items 6, 7, 16, 19, and 20; Loading is the standardized value of the item loadings.

Table 4.5. Factor Loadings from Grade-3 Test CFA

Bank item number	Final-scale item #	Loading	s.e.
WP26.C	1	0.65	0.05
WP16.C	2	0.72	0.04
WP35.C	3	0.62	0.04
WP43.C	4	0.68	0.03
WP55.C	5	0.52	0.04
ES15.S1	6	0.58	0.08
ES1.3.S1	7	0.57	0.05
ES18.S1	8	0.73	0.03
FPW3.5	9	0.53	0.04
FPW14.S1	10	0.43	0.05
FNL2.1.S1	11	0.52	0.04
FWP7.S1	12	0.57	0.04
FWP7.C	13	0.59	0.04
CWN4.1.C	14	0.75	0.05
CWN3.4.C	15	0.69	0.05
ES3.3.C	16	0.90	0.02
CTWN6.1	17	0.51	0.05
CWN42.C	18	0.67	0.04
CWN49.C	19	0.64	0.04
ES.27.C	20	0.87	0.02
CF7.C	21	0.45	0.04
ES21.C	22	0.88	0.02
CWN33.2	23	0.51	0.04
CWN50.C	24	0.65	0.03

Note. Final-scale item # = the newly generated item number after excluding items 6, 7, 16, and 19; Loading is the standardized value of the item loadings.

5. Classical Test Theory Analyses

We conducted analyses based on classical test theory (CTT) after the findings of the dimensionality of the test at each grade level were confirmed.

5.1. Distribution of the Observed Test Score

Figures 5.1–5.4 illustrate the minimum and maximum of the total test score for each grade level as well as the percentages of students at the observed test score. In the sample of 568 kindergarteners, 31 received a perfect score of 15, and none of the students received the minimum score of 0. The mean and standard deviation of the distribution were 9.75 and 3.22, respectively.

Figure 5.1. Bar graph depicting the distribution of the observed test scores for the kindergarten test.

Figure 5.2. Bar graph depicting the distribution of the observed test scores for the grade-1 test.

The grade-1 test included 18 dichotomously scored items, and the minimum and maximum possible observed score of the final test form were 0 and 18, respectively. In the sample of 636 grade-1 students, 6 received perfect scores, and 1 received the minimum score of 0. The mean and standard deviation are 11.25 and 3.60, respectively.

Figure 5.3. Bar graph depicting the distribution of the observed test scores for the grade-2 test.

The grade-2 test included 20 dichotomously scored items, and the minimum and maximum possible observed score of the final test form were 0 and 20, respectively. In the sample of 650 grade-2 students, 9 students received perfect scores, and 1 received the minimum score of 0. The mean and standard deviation are 12.30 and 4.11, respectively.

Figure 5.4. Bar graph depicting the distribution of the observed test scores for the grade-3 test.

The grade-3 test included 24 dichotomously scored items, so the minimum and maximum possible observed score of the final test form were 0 and 24, respectively. In the sample of 684 grader-3 students, 24 received perfect scores, and 0 received the minimum score of 0. The mean and standard deviation are 16.11 and 5.30, respectively.

5.2. Item Difficulty and Discrimination

Item difficulty (P-value) and item discrimination estimates were calculated on the basis of the CTT approach. Item difficulty estimates were calculated as the proportions of correct answers, which were also equal to the item means. Item discrimination estimates were calculated as the item-rest correlation coefficients. Tables 5.1–5.4 illustrate the descriptive statistics, item difficulty, and item discrimination estimates of the items of each grade-level test; Table 5.5 shows the information on distribution of P-value and item-rest correlations for the test at each grade level.

Table 5.1. Item Difficulty and Discrimination from classical test theory (CTT) Analyses of Kindergarten Test

Bank item number	Final-scale item #	Mean	St. dev.	Item-rest <i>r</i>
CT1.4.S1	1	0.95	0.22	0.17
WP22.S1	2	0.90	0.31	0.31
WP34.S2	3	0.46	0.50	0.32
WP24.C	4	0.48	0.50	0.51
ES15.S1	5	0.60	0.49	0.37
ES1.3.S1	6	0.55	0.50	0.17
FPW2.2.S1	7	0.45	0.50	0.40
FPW3.3.S1	8	0.30	0.46	0.29
CBF15.S2	9	0.90	0.29	0.33
CBF21.S2	10	0.71	0.46	0.54
CBF3.S2	11	0.70	0.46	0.51
CWN1.1.S2	12	0.88	0.33	0.35
CWN2.1.S2	13	0.73	0.45	0.43
CBF35.C	14	0.59	0.49	0.46
CWN25.C	15	0.56	0.50	0.50

Note. Final-scale item # = the newly generated item number after excluding items 5 and 6.

Table 5.2. Item Difficulty and Discrimination from CTT Analyses of Grade-1 Test

Bank item number	Final-scale item #	Mean	St. dev.	Item-rest <i>r</i>
WP34.S2	1	0.61	0.49	0.49
WP24.C	2	0.75	0.43	0.52
WP4.2.C	3	0.75	0.43	0.44
WP43.C	4	0.31	0.46	0.46
ES15.S1	5	0.83	0.38	0.32
ES1.3.S1	6	0.72	0.45	0.35
ES18.S1	7	0.54	0.50	0.35
FPW2.2.S1	8	0.80	0.40	0.47
FPW3.3.S1	9	0.42	0.49	0.33
FPW3.5.S1	10	0.26	0.44	0.30
FPW14.S1	11	0.33	0.47	0.20
CBF3.S2	12	0.91	0.28	0.25
CWN2.1.S2	13	0.87	0.34	0.34
CWN4.1.C	14	0.66	0.47	0.51
CBF35.C	15	0.86	0.34	0.27
ES3.3.C	16	0.24	0.42	0.37
CWN25.C	17	0.87	0.34	0.33
CTWN6.1.C	18	0.53	0.50	0.29

Note. Final-scale item # = the newly generated item number after excluding items 5, 6, 15, and 19.

Table 5.3. Item Difficulty and Discrimination from CTT Analyses of Grade-2 Test

Bank item number	Final-scale item #	Mean	St. dev.	Item-rest <i>r</i>
WP4.2.C	1	0.83	0.38	0.42
WP26.C	2	0.63	0.48	0.44
WP16.C	3	0.61	0.49	0.48
WP43.C	4	0.53	0.50	0.49
WP55.C	5	0.22	0.42	0.35
ES15.S1	6	0.90	0.30	0.30
ES1.3.S1	7	0.83	0.38	0.37
ES18.S1	8	0.60	0.49	0.51
FPW2.2.S1	9	0.87	0.34	0.34
FPW3.3.S1	10	0.56	0.50	0.29
FPW3.5.S1	11	0.25	0.43	0.25
FPW14.S1	12	0.43	0.50	0.19
CBF3.S2	13	0.95	0.21	0.25
CWN4.1.C	14	0.88	0.33	0.43
CWN3.4.C	15	0.80	0.40	0.40
ES3.3.C	16	0.37	0.48	0.52
CTWN6.1.C	17	0.64	0.48	0.37
CWN42.C	18	0.56	0.50	0.48
ES21.C	19	0.27	0.44	0.54
CWN33.2.C	20	0.56	0.50	0.32

Note. Final-scale item # = the newly generated item number after excluding items 6, 7, 16, 19, and 20.

Table 5.4. Item Difficulty and Discrimination from CTT Analyses of Grade-3 Test

Bank item number	Final-scale item #	Mean	St. dev.	Item-rest <i>r</i>
WP26.C	1	0.91	0.29	0.37
WP16.C	2	0.78	0.42	0.52
WP35.C	3	0.57	0.49	0.47
WP43.C	4	0.61	0.49	0.51
WP55.C	5	0.26	0.44	0.34
ES15.S1	6	0.96	0.19	0.26
ES1.3.S1	7	0.89	0.32	0.34
ES18.S1	8	0.72	0.45	0.51
FPW3.5.S1	9	0.56	0.50	0.40
FPW14.S1	10	0.57	0.50	0.32
FNL2.1.S1	11	0.68	0.47	0.39
FWP7.S1	12	0.69	0.46	0.43
FWP7.C	13	0.75	0.43	0.43
CWN4.1.C	14	0.90	0.30	0.43
CWN3.4.C	15	0.86	0.35	0.43
ES3.3.C	16	0.51	0.50	0.60
CTWN6.1.C	17	0.77	0.42	0.35
CWN42.C	18	0.77	0.42	0.49
CWN49.C	19	0.48	0.50	0.48
ES.27.C	20	0.65	0.48	0.64
CF7.C	21	0.61	0.49	0.34
ES21.C	22	0.47	0.50	0.58
CWN33.2.C	23	0.68	0.47	0.38
CWN50.C	24	0.48	0.50	0.48

Note. Final-scale item # = the newly generated item number after excluding items 6, 7, 16, and 19.

Table 5.5. Descriptive Statistics of CTT-based Item Parameter Estimates of the Test at Each Grade Level

Value	Number of items			
	Kindergarten	Grade 1	Grade 2	Grade 3
<i>Item difficulty</i>				
>0.90	3	1	2	3
0.80–0.89	1	4	5	2
0.70–0.79	3	4	0	5
0.60–0.69	1	2	4	6
0.50–0.59	3	2	4	4
0.40–0.49	3	1	1	3
0.30–0.39	1	2	1	0
0.20–0.29	0	2	3	1
0.10–0.19	0	0	0	0
<0.09	0	0	0	0
Mean	0.65	0.63	0.62	0.67
Median	0.60	0.69	0.61	0.68
St. dev.	0.20	0.23	0.23	0.17
<i>Item-rest r</i>				
0.80–1.00	0	0	0	0
0.60–0.79	0	0	0	2
0.40–0.59	7	6	10	13
0.20–0.39	6	12	9	9
0.00–0.20	2	0	1	0
Mean	0.38	0.37	0.39	0.44
Median	0.37	0.35	0.38	0.43
St. dev.	0.12	0.09	0.10	0.09

5.3. Coefficients α and Standard Error of Measurement

Table 5.6 shows the coefficient α (Cronbach, 1951) of the test at each grade level. We calculated the standard error of measurement (SEM) of the test at each grade level. The scale variances are also reported in Table 5.6. The SEM was calculated as in Equation 1:

$$SEM = \sqrt{\sigma^2 \times (1 - \rho_{XX})} \quad (1)$$

Table 5.6. Coefficient α of the Test at Each Grade Level

Grade level	Number of items	σ^2_{scale}	Coefficient α	SEM
Kindergarten	15	10.40	0.77	1.55
1	18	12.98	0.79	1.65
2	20	16.92	0.82	1.75
3	24	28.06	0.87	1.91

Note. SEM = standard error of measurement; σ^2_{scale} = scale variances.

6. Item Response Theory (IRT) Analyses

6.1. Model Description

The IRT analyses were carried out by flexMIRT 3.5 (Cai, 2017). All items in each grade level were dichotomously scored, both the constructed-response items and the selected-response items. Although adjustment for the guessing effect on such items can be done in the three-parameter logistic model, de Ayala (2009) suggests at least a sample size of 1000 to have stable estimates in IRT calibration. Because our sample sizes in each grade level were less than 1000, the two-parameter logistic (2PL) model was found to be an appropriate model to fit the data. The formula of the 2PL model is shown in Equation 2 according to the parameterization of de Ayala (2009).

$$P_j(\theta) = \frac{\exp [a_j(\theta - b_j)]}{1 + \exp [a_j(\theta - b_j)]} \quad (2)$$

where a_j is the discrimination index of item j ($j = 1, 2, \dots, J$), b_j is the difficulty index of item j , P_j is the probability of correct answer, θ is the person ability.

After the most appropriate IRT model was chosen, vertical scaling was implemented by means of the fixed item parameter calibration (FIPC) approach under the common item nonequivalent group test design so that we could place scores from the tests of all grade levels on a common scale. In doing so, we get produced scores comparable across grades. According to Kolen and Brennan (2004), vertical scaling is conducted for tests that have some differences in difficulty but that are developed to measure the same or similar constructs. The common item nonequivalent group test design is one of the data-collection designs for conducting the vertical scaling. In this design, adjacent grade levels share common or anchor items. For example, an item block from the grade-1 test form can be also administered to the grade-2 students. The purpose of the FIPC is to use these items common to adjacent grades (e.g., grade 1 and grade 2) or any two grades (e.g., grade 1 and grade 3) to estimate the noncommon items. First a base scale must be chosen. We chose the grade 1 as the base scale, so kindergarten, grade 2, and grade 3 are linked to the grade 1. Table 6.1 shows the number of items common to the grades. The results of flexMIRT of all grade levels indicated successful convergences.

Below, both latent trait and item parameter estimates for each grade level on a common scale, i.e., vertically scaled estimates, are presented. We also report the within-grade item parameter estimates and latent trait estimates for each grade level, in Appendix J. In Appendix K, we report how the θ scores for additional data from EMSA 2018 K–3 tests were scaled.

Table 6.1. The Number of Items Common to Grade Levels

Grade level	K	1	2	3
K	15			
1	10	18		
2	5	13	20	
3	2	9	16	24

Note. The numbers of items presented to each grade are on the diagonal. Items common to all grade levels are on the off diagonals. Items common to grade 1 and the other grades appear in boldface type.

6.2. Item Difficulty and Discrimination

Table 6.2 provides the descriptive statistics of the item parameter estimates of each item in each grade level test. In addition, Tables 6.3–6.6 show the parameter estimates in each grade-level test, and Figures 6.1–6.8 demonstrate the distributions of the parameter estimates in each grade-level test. The means of the discrimination parameter estimates across grade levels were like each other while the means of the discrimination parameter estimates showed an increasing pattern from kindergarten to grade 3. In addition, the standard deviations of the difficulty-parameter estimates seemed to be similar across grade levels except for grade 3, whereas the standard deviations of the discrimination-parameter estimates were similar across grade levels except for grade 2. The easiest item of the test had values of -4.57 , -2.57 , -2.57 , and -1.79 for kindergarten, grade 1, grade 2, and grade 3, respectively, whereas the most difficult had values of 0.45 , 1.47 , 1.96 , and 2.87 for kindergarten, grade 1, grade 2, and grade 3, respectively. Moreover, the discrimination-parameter estimates of each grade level test were 0.5 or higher.

Table 6.2. Descriptive Statistics of Discrimination Estimates and Difficulty Estimates of Each Item in Each Grade Level Test

	Kindergarten		Grade 1		Grade 2		Grade 3	
	<i>a</i>	<i>b</i>	<i>a</i>	<i>b</i>	<i>a</i>	<i>b</i>	<i>a</i>	<i>b</i>
Mean	1.42	-1.89	1.28	-0.59	1.30	0.00	1.17	0.39
St. dev.	0.43	1.22	0.46	1.23	0.56	1.17	0.42	1.05
Minimum	0.85	-4.57	0.49	-2.57	0.49	-2.57	0.49	-1.79
Maximum	2.27	0.45	2.27	1.47	3.09	1.96	2.21	2.87
Skewness	0.49	-0.21	0.39	0.22	1.72	-0.35	0.90	0.09
Kurtosis	-0.71	0.77	-0.39	-0.95	4.82	-0.24	0.50	0.30

Note. *a* = item discrimination index; *b* = item difficulty index.

Table 6.3. Parameter Estimates and Standard Errors for Final-Scale Items of Kindergarten Test Modeled by two-parameter logistic (2PL) with fixed item parameter calibration (FIPC) Approach.

Bank item number	Final-scale item #	a	$s.e.$	b	$s.e.$
CT1.4.S1	1	0.95	0.27	-4.57	0.81
WP22.S1	2	1.45	0.23	-3.02	0.21
WP34.S2	3	1.74	----	-0.39	----
WP24.C	4	2.27	----	-0.85	----
ES15.S1	5	1.07	----	-1.79	----
ES1.3.S1	6	1.03	----	-1.10	----
FPW2.2.S1	7	1.78	----	-1.15	----
FPW3.3.S1	8	0.85	----	0.45	----
CBF15.S2	9	1.69	0.26	-2.94	0.20
CBF21.S2	10	2.07	0.24	-1.79	0.07
CBF3.S2	11	1.08	----	-2.57	----
CWN1.1.S2	12	1.59	0.25	-2.77	0.18
CWN2.1.S2	13	1.42	----	-1.75	----
CBF35.C	14	0.96	----	-2.23	----
CWN25.C	15	1.30	----	-1.87	----

Note. Final-scale item # = the newly generated item number after excluding items 5 and 6.

---- Standard errors were not calculated, because item parameters were fixed.

Table 6.4. Parameter Estimates and Standard Errors for Final-Scale Items of Grade-1 Test Modeled by 2PL.

Bank item number	Final-scale item #	<i>a</i>	<i>s.e.</i>	<i>b</i>	<i>s.e.</i>
WP34.S2	1	1.74	0.20	–0.39	0.08
WP24.C	2	2.27	0.27	–0.85	0.08
WP4.2.C	3	1.57	0.19	–1.01	0.11
WP43.C	4	1.83	0.23	0.70	0.08
ES15.S1	5	1.07	0.18	–1.79	0.24
ES1.3.S1	6	1.03	0.15	–1.10	0.15
ES18.S1	7	0.92	0.13	–0.21	0.11
FPW2.2.S1	8	1.78	0.21	–1.15	0.11
FPW3.3.S1	9	0.85	0.12	0.45	0.12
FPW3.5.S1	10	0.97	0.15	1.27	0.18
FPW14.S1	11	0.49	0.11	1.47	0.34
CBF3.S2	12	1.08	0.20	–2.57	0.38
CWN2.1.S2	13	1.42	0.22	–1.75	0.19
CWN4.1.C	14	1.73	0.20	–0.59	0.08
CBF35.C	15	0.96	0.17	–2.23	0.32
ES3.3.C	16	1.32	0.16	1.17	0.13
CWN25.C	17	1.30	0.21	–1.86	0.22
CTWN6.1.C	18	0.73	0.12	–0.17	0.13

Note. Final-scale item # = the newly generated item number after excluding items 5, 6, 15, and 19.

Table 6.5. Parameter Estimates and Standard Errors for Final-Scale Items of Grade-2 Test Modeled by 2PL with FIPC Approach.

Bank item number	Final-scale item #	<i>a</i>	<i>s.e.</i>	<i>b</i>	<i>s.e.</i>
WP4.2.C	1	1.57	----	-1.01	----
WP26.C	2	1.24	0.14	0.09	0.09
WP16.C	3	1.50	0.18	0.21	0.08
WP43.C	4	1.83	----	0.70	----
WP55.C	5	1.20	0.15	1.96	0.14
ES15.S1	6	1.07	----	-1.79	----
ES1.3.S1	7	1.03	----	-1.10	----
ES18.S1	8	0.92	----	-0.21	----
FPW2.2.S1	9	1.78	----	-1.15	----
FPW3.3.S1	10	0.85	----	0.45	----
FPW3.5.S1	11	0.97	----	1.27	----
FPW14.S1	12	0.49	----	1.47	----
CBF3.S2	13	1.08	----	-2.57	----
CWN4.1.C	14	1.73	----	-0.59	----
CWN3.4.C	15	1.31	0.17	-0.75	0.15
ES3.3.C	16	1.32	----	1.17	----
CTWN6.1.C	17	0.73	----	-0.17	----
CWN42.C	18	1.42	0.16	0.39	0.07
ES21.C	19	3.09	0.39	1.35	0.07
CWN33.2.C	20	0.79	0.11	0.26	0.12

Note. Final-scale item # = the newly generated item number after excluding items 6, 7, 16, 19, and 20.

---- Standard errors were not calculated because item parameters were fixed.

Table 6.6. Parameter Estimates and Standard Errors for Final-Scale Items of Grade-3 Test Modeled by 2PL with FIPC Approach.

Bank item number	Final-scale item #	a	$s.e.$	b	$s.e.$
WP26.C	1	1.29	0.18	-1.11	0.23
WP16.C	2	1.48	0.16	0.05	0.11
WP35.C	3	1.11	0.13	0.98	0.10
WP43.C	4	1.83	----	0.70	----
WP55.C	5	0.86	0.12	2.87	0.20
ES15.S1	6	1.07	----	-1.79	----
ES1.3.S1	7	1.03	----	-1.10	----
ES18.S1	8	0.92	----	-0.21	----
FPW3.5.S1	9	0.97	----	1.27	----
FPW14.S1	10	0.49	----	1.47	----
FNL2.1.S1	11	0.81	0.10	0.21	0.16
FWP7.S1	12	0.94	0.11	0.28	0.14
FWP7.C	13	1.01	0.12	-0.05	0.16
CWN4.1.C	14	1.73	----	-0.59	----
CWN3.4.C	15	1.31	0.16	-0.58	0.19
ES3.3.C	16	1.32	----	1.17	----
CTWN6.1.C	17	0.73	----	-0.17	----
CWN42.C	18	1.27	0.16	0.00	0.13
CWN49.C	19	1.15	0.13	1.45	0.09
ES.27.C	20	2.21	0.23	0.78	0.07
CF7.C	21	0.68	0.09	0.58	0.15
ES21.C	22	1.97	0.19	1.47	0.07
CWN33.2.C	23	0.78	0.10	0.23	0.16
CWN50.C	24	1.15	0.13	1.47	0.09

Note. Final-scale item # = the newly generated item number after excluding items 6, 7, 16, and 19.

---- Standard errors were not calculated because item parameters were fixed.

Figure 6.1. Item discrimination estimate (a) of each final-scale item of kindergarten test.

Figure 6.2. Item difficulty estimate (b) of each final-scale item of kindergarten test.

Figure 6.3. Item discrimination estimate (a) of each final-scale item of grade-1 test.

Figure 6.4. Item difficulty estimate (b) of each final-scale item of grade-1 test.

Figure 6.5. Item discrimination estimate (a) of each final-scale item of grade-2 test.

Figure 6.6. Item difficulty estimate (b) of each final-scale item of grade-2 test.

Figure 6.7. Item discrimination estimate (a) of each final-scale item of grade-3 test.

Figure 6.8. Item difficulty estimate (b) of each final-scale item of grade-3 test.

6.3. Test Information Function

Test information is the amount of information that the characteristics of the test provide for person location estimates. de Ayala (2019) recommended Equation 3, describing the relationship between the test information and conditional standard error of measurement (CSEM), where I represents the test information function for a given person ability (θ).

$$CSEM(\theta) = \frac{1}{\sqrt{I(\theta)}} \quad (3)$$

The test information values and CSEMs based on a common scale (i.e., based on the parameter estimates from the vertical scaling implementation) were calculated by means of flexMIRT software.

Figure 6.9 shows this relationship for the kindergarten student test. The figure shows that the person ability estimates around -1.20 were associated with the highest test information and the lowest CSEM. Furthermore, low CSEM values (i.e., low accuracy of person-ability estimation) appeared when person ability estimates were lower than -3.00 and higher than 0.40 . Approximately 94% of the kindergarten students fell within the interval between these two values on the vertical scale.

Figure 6.9. Test information curve and conditional standard error of measurement (CSEM) for the final scale format of kindergarten test.

Figure 6.10 shows that the person-ability estimates around -0.60 were associated with the highest test information and the lowest CSEM in the grade-1 test, whereas low CSEM values appeared when person-ability estimates were lower than -2.00 and higher than 1.00 . Approximately 84% of the grade-1 students fell within the interval between these two values on the vertical scale.

Figure 6.10. Test information curve and CSEM for the final scale format of-grade-1 test.

For the grade-2 test shown in Figure 6.11, person ability estimates around 1.00 were associated with the highest test information and the lowest CSEM. In addition, low CSEM values appeared when person ability estimates were lower than -1.60 and higher than 2.00 . Approximately 91% of the grade-2

students fell within the interval between those two values on the vertical scale.

Figure 6.11. Test information curve and CSEM for the final scale format of grade-2 test.

Figure 6.12 shows that person-ability estimates around 1.00 were associated with the highest test information and the lowest CSEM in the grade-3 test, whereas high CSEM values appeared when person ability estimates were lower than -1.20 and higher than 2.20. Approximately 74% of the grade-3 students fell within the interval between those two values on the scale that was equated across grade levels.

Figure 6.12. Test information curve and CSEM for the final scale format of grade-3 test.

6.4. Latent Trait Estimation

Maximum likelihood estimation (MLE) was used to estimate the latent person ability at each grade level. Note that these estimates are not available for those with perfect scores. We also used the expected *a posteriori* method to estimate the person ability at each grade level. Figures 6.13–6.20 show the distributions of these two types of person ability estimates based on a common scale.

Figure 6.13. Maximum likelihood estimate (MLE)-based latent estimates of kindergarten students.

Figure 6.14. Expected a posteriori (EAP)-based latent estimates of kindergarten students.

Figure 6.15. MLE-based latent estimates of grade-1 students.

Figure 6.16. EAP-based latent estimates of grade-1 students.

Figure 6.17. MLE-based latent estimates of grade-2 students.

Figure 6.18. EAP-based latent estimates of grade-2 students.

Figure 6.19. MLE-based latent estimates of grade-3 students.

Figure 6.20. EAP-based latent estimates of grade-3 students.

References

- Cai, L. (2017). flexMIRT: Flexible Multilevel Multidimensional Item Analysis and Test Scoring (Version 3.51) [Computer Software]. Vector Psychometric Group.
- Carpenter, T. P., Ansell, E., Franke, M. L., Fennema, E., & Weisbeck, L. (1993). Models of problem solving: A study of kindergarten children's problem-solving processes. *Journal for Research in Mathematics Education*, 24(5), 428–441.
- Carpenter, T. P., Fennema, E., Franke, M. L., Levi, L., & Empson, S. B. (1999). *Children's Mathematics: Cognitively Guided Instruction*. Heinemann.
- Carpenter, T. P., Fennema, E., Franke, M. L., Levi, L., & Empson, S. B. (2015). *Children's Mathematics: Cognitively Guided Instruction* (2nd Ed.). Heinemann.
- Carpenter, T. P., Franke, M. L., & Levi, L. (2003). *Thinking Mathematically: Integrating Arithmetic and Algebra in Elementary School*. Heinemann.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297–334. <https://doi.org/10.1007/BF02310555>
- de Ayala, R. J. (2009). *The Theory and Practice of Item Response Theory*. The Guilford Press.
- Dixon, J. K., Larson, M., Leiva, M. A., & Adams, T. L. (2013). *GoMath! Florida*. Orlando, FL: Houghton Mifflin Harcourt.
- FileMaker Pro (Version 16) [Computer Software]. Santa Clara, CA: Filemaker, Inc.
- Flake, J. K., Pek, J., & Hehman, E. (2017). Construct validation in social and personality research: Current practice and recommendations. *Social Psychological and Personality Science*, 8(4), 370–378.
- Glorfeld (1995). An improvement on Horn's parallel analysis methodology for selecting the correct number of factors to retain. *Educational and Psychological Measurement*, 55(3), 377–393. <https://doi.org/10.1177/0013164495055003002>
- Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, 30(2), 179–185. <https://doi.org/10.1007/BF02289447>
- IBM Corp. (2017). *IBM SPSS Statistics for Windows, Version 25.0*. IBM Corp.
- Kolen, M. J., & Brennan, R. L. (2014). *Test equating, scaling, and linking: Methods and practices*. Springer Science & Business Media.
- Muthén, L. K., & Muthén, B. O. (2017). *Mplus: Statistical Analysis with Latent Variables: User's Guide* (Version 8). Authors.
- Muthén, B. O., du Toit, S. H. C., & Spisic, D. (1997). Robust interference using weighted least squares and quadratic estimating equations in the latent variable modeling with categorical and continuous outcomes. *Unpublished manuscript, University of California, Los Angeles, USA*.
- NGACBP & CCSSO (National Governors Association Center for Best Practices & Council of Chief State School Officers) (2010). *Common Core State Standards for Mathematics*. Washington, DC: National Governors Association Center for Best Practices & Council of Chief State School Officers.

- Pothier, Y., & Sawada, D. (1983). Partitioning: The emergence of rational number ideas in children. *Journal for Research in Mathematics Education*, 14(4), 307–317.
- R Core Team (2020). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>
- Revelle, W. (2020). *psych: Procedures for Psychological, Psychometric, and Personality Research*. Northwestern University, Evanston, Illinois. R package version 1.9.12. <https://CRAN.Rproject.org/package=psych>
- Schoen, R. C., LaVenía, M., Bauduin, C., & Farina, K. (2016a). Elementary mathematics student assessment: Measuring the performance of grade 1 and 2 students in counting, word problems, and computation in fall 2013 (Research Report No. 2016-03). Florida State University. <http://dx.doi.org/10.17125/fsu.1508170543>
- Schoen, R. C., LaVenía, M., Bauduin, C., & Farina, K. (2016b). Elementary mathematics student assessment: Measuring the performance of grade 1 and 2 students in counting, word problems, and computation in fall 2014 (Research Report No. 2016-04). Florida State University. <http://dx.doi.org/10.17125/fsu.1508174887>
- Schoen, R. C., LaVenía, M., Champagne, Z. M., & Farina, K., (2016c). Mathematics performance and cognition (MPAC) interview: Measuring first- and second-grade student achievement in number, operations, and equality in spring 2014 (Research Report No. 2016–01). Florida State University. <http://dx.doi.org/10.17125/fsu.1493238156>
- Schoen, R. C., LaVenía, M., Champagne, Z. M., Farina, K., & Tazaz, A. (2016d). Mathematics performance and cognition (MPAC) interview: Measuring first- and second-grade student achievement in number, operations, and equality in spring 2015 (Research Report No. 2016–02). Florida State University. <http://dx.doi.org/10.17125/fsu.1493238666>
- Schoen, R. C., Anderson, D., Champagne, Z., & Bauduin, C. (2017). Elementary mathematics student assessment: Measuring the performance of grade K, 1, and 2 students in counting, word problems, and computation in fall 2015. (Research Report No. 2017-20). Florida State University. <http://dx.doi.org/10.17125/fsu.1522170756>
- Schoen, R. C., Liu, S., Yang, X., & Paek, I. (2017). *Psychometric report for the Early Fractions Test administered with third- and fourth-grade students in fall 2016* (Research Report No. 2017-10). Florida State University. <https://fsu.digital.flvc.org/islandora/object/fsu%3A535846>
- Schoen, R. C., Anderson, D., Riddell, C. M., & Bauduin, C. (2018). Elementary mathematics student assessment: Measuring the performance of grade 3, 4, and 5 students in number (whole numbers and fractions), operations, and algebraic thinking in fall 2015 (Research Report No. 2018-24). Florida State University.
- Schoen, R. C., Buntin, C., Guven, A., & Yang, X. (2021a). Elementary mathematics student assessment: Measuring the performance of grade K, 1, 2, and 3 students in number (whole numbers and fractions), operations, and algebraic thinking in fall 2018. (Research Report No. 2021-05). Florida State University.
- Schoen, R. C., Champagne, Z. M., Whitacre, I., & McCrackin, S. (2021b). Comparing the number and variation of additive word problems in first-grade U.S. textbooks in the 1980s and the Common Core era. *School Science and Mathematics*, 121(2), 110–121. <https://doi.org/10.1111/ssm.12447>

- Turner, E. E., & Celedón-Pattichis, S. (2011). Mathematical problem solving among Latina/o kindergartners: An analysis of opportunities to learn. *Journal of Latinos and Education, 10*(2), 146–169.
- Verschaffel, L., Greer, B., & De Corte, E. (2007). Whole numbers concepts and operations. In F. K. Lester, Jr. (Ed.), *Second Handbook of Research on Mathematics Teaching and Learning*. National Council of Teachers of Mathematics.
- Zwick, W. R., & Velicer, W. F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin, 99*(3), 432–442. <https://doi.org/10.1037/0033-2909.99.3.432>

Appendix A. Grade K Test

The form in this appendix is identical to the form used in spring 2019. As a result, there are no headers or footers present in this section of the report.

Kindergarten – Spring Elementary Mathematics Student Assessment

Date: _____

District: _____

School: _____

Teacher: _____

Student: _____

Grade: _____

This test may include some kinds of problems that are new or challenging for you. Don't worry if you can't solve them. You won't be graded on this test, but please try your best!

This page is intentionally left blank.

Example response:

What grade are you in?

K

1

2

3

☒

☐

☐

☐

Draw a circle around the picture of the car.

Write the number four in the box.

4

0

0

0

0

0

0

0

0

0

0

0

0

Yes

No

Yes

No

Yes

No

Yes

No

=

0

0

0

0

=

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

Appendix B. Grade 1 Test

The form in this appendix is identical to the form used in spring 2019. As a result, there are no headers or footers present in this section of the report.

First Grade – Spring Elementary Mathematics Student Assessment

Date: _____

District: _____

School: _____

Teacher: _____

Student: _____

Grade: _____

This test may include some kinds of problems that are new or challenging for you. Don't worry if you can't solve them. You won't be graded on this test, but please try your best!

This page is intentionally left blank.

Example responses:

What grade are you in?

1

2

3

4

0

0

0

0

Draw a circle around the picture of the car.

Write the number four in the box.

4

0

0

0

0

Yes

No

Yes

No

Yes

No

Yes

No

Yes

No

0

0

0

=

0

0

0

0

=

0

0

0

0

0

0

0

0

=

=

Appendix C. Grade 2 Test

The form in this appendix is identical to the form used in spring 2019. As a result, there are no headers or footers present in this section of the report.

Second Grade – Spring Elementary Mathematics Student Assessment

Date: _____

District: _____

School: _____

Teacher: _____

Student: _____

Grade: _____

This test may include some kinds of problems that are new or challenging for you. Don't worry if you can't solve them. You won't be graded on this test, but please try your best!

This page is intentionally left blank.

Example responses:

What grade are you in?

1	2	3	4
0	0	0	0

Draw a circle around the picture of the car.

Write the number four in the box.

Yes

No

Yes

No

Yes

No

Yes

No

Yes

No

0

0

0

0

0

0

0

0

0

0

0

$$\begin{array}{|c|} \hline \\ \hline \end{array} + \begin{array}{|c|c|} \hline \\ \hline \end{array} = \begin{array}{|c|} \hline \\ \hline \end{array}$$

$$\begin{array}{|c|c|} \hline \\ \hline \end{array} + \begin{array}{|c|c|} \hline \\ \hline \end{array} = \begin{array}{|c|} \hline \\ \hline \end{array}$$

$$\begin{array}{|c|} \hline \\ \hline \end{array} + \begin{array}{|c|} \hline \\ \hline \end{array} + \begin{array}{|c|} \hline \\ \hline \end{array} = \begin{array}{|c|} \hline \\ \hline \end{array} + \begin{array}{|c|} \hline \\ \hline \end{array}$$

Appendix D. Grade 3 Test

The form in this appendix is identical to the form used in spring 2019. As a result, there are no headers or footers present in this section of the report.

Third Grade – Spring Elementary Mathematics Student Assessment

Date: _____

District: _____

School: _____

Teacher: _____

Student: _____

Grade: _____

This test may include some kinds of problems that are new or challenging for you. Don't worry if you can't solve them. You won't be graded on this test, but please try your best!

This page is intentionally left blank.

Example responses:

What grade are you in?

1

2

3

4

0

0

0

0

Draw a circle around the picture of the car.

Write the number four in the box.

4

Yes

No

Yes

No

Yes

No

Yes

No

Yes

No

0

0

0

0

0

0

0

0

0

0

0

Answer: _____

0

0

0

0

Appendix E. Grade K Administration Guide

The form in this appendix is identical to the form used in spring 2019. As a result, there are no headers or footers present in this section of the report. The EMSA test was administered with two other assessments and those assessments are reflected in the contents of the administration guide.

Administration Instructions for the REALM Project Kindergarten Spring Assessments and Surveys

2018–2019

Teacher Name: _____

EMSA Tests (Blue Booklet)

Administration Date: _____

Start Time: _____ **End Time:** _____

Student Surveys (Green Booklet)

Administration Date: _____

Start Time: _____ **End Time:** _____

This page is intentionally left blank.

Table of Contents

Administration Guidelines for Assessments and Student Surveys	1
Overview	1
Materials	1
EMSA Test, Surveys, and Reading Booklets.....	1
Students to be Tested	2
Preparing for Testing	2
Administering the EMSA, Student Surveys, and Reading Test	3
Accommodations	4
Testing Time Allocation	4
Submitting Materials	4
Spring 2019 EMSA Administration Instructions.....	5

This page is intentionally left blank.

Administration Guidelines for Assessments and Student Surveys

Overview

Thank you for your participation in the REALM Project. This document will provide you with instructions to follow for the purpose of assessing your mathematics students. Within this document are administration guidelines and instructions for assessments and student surveys. There is also important information about the testing window, as well as returning all materials to the REALM Project staff.

The assessments and student surveys are designed to be administered in a written format with the whole class, but you may choose to administer individually or in small groups as you see fit. Please administer the Spring 2019 Elementary Mathematics Student Assessment (EMSA), student surveys, and reading test during the allotted testing window of April 15, 2019 to May 3, 2019. If you cannot administer the assessments during that window, please notify Charity Bauduin (cbauduin2@lsi.fsu.edu) and plan to administer them as soon as possible. You may choose to administer the EMSA test, student surveys, and reading test on separate days. If you choose to administer all assessments and surveys in the same day, you may consider providing students with a brief stretch break in between each assessment. We do ask that you administer the EMSA first, followed by the student surveys, and conclude with the reading test.

Scripts for the teacher to use during administration begin on page 7 of this guide. Please follow the scripts as closely as possible when you or your surrogate administer the assessments and surveys. The assessments, surveys, and administration guides will be picked up as described in the Submitting the Materials section on page 4.

We ask that students use pencils to complete these assessments.

Materials

The following materials are required for testing:

- Administration Instructions for the Kindergarten Spring Assessments and Surveys (this document)
- An EMSA booklet for each student (blue packet, one per student, provided)
- A Student Surveys booklet for each student (green packet, one per student, provided)
- A FSU REALM Project pencil for each student (one per student, provided)
- At least one sharpened pencil for each student

EMSA Test, Surveys, and Reading Booklets

Students should mark their answers directly in the booklets. Should you need additional testing materials, please contact Charity Bauduin (cbauduin2@lsi.fsu.edu). Remember that these materials are to remain at the school site until the testing window has ended. The materials should be stored in a secure, access-restricted location at all times.

Students to be Tested

We ask that you administer the math assessment (EMSA), student surveys, and reading test to students for whom you are the teacher of record. If you teach multiple groups of students mathematics, please contact Charity Bauduin (cbauduin2@lsi.fsu.edu) for additional testing materials. We have provided enough materials so that all students may complete the assessments and surveys. However, we will only enter and analyze data for those children who meet the requirements for consent.

Preparing for Testing

The first page of each EMSA booklet (blue packet) has the following box for student information:

Date:	
District:	School
Teacher:	
Student:	Grade

We have attached labels for those students who were included on the class rosters provided to us. If you have students with pre-labeled booklets that have moved, please cross out their names. Please use these booklets for any student who does not have pre-labeled booklets. There are additional booklets containing no labels that can also be used for students who do not have pre-labeled booklets. For those students who do not have pre-labeled EMSA tests, we ask that the classroom teacher, or an adult surrogate, enter this information (date, district name, school name, teacher name, student full name as it appears on official records, and student grade level) on each test booklet for each student to be tested prior to the testing session. (Please do not leave it for students to enter this information.)

The first page of each Student Surveys booklet (green packet) has the following box for student information:

Student’s Name:	Date:
Teacher:	

Prior to the testing session, the classroom teacher, or an adult surrogate, must enter this information (student's name full name as it appears on official records, date, and teacher name) on each survey booklet for each student to be tested who does not have a pre-labeled booklet. (Please do not leave it for students to enter this information.)

The Spring EMSA, Student Surveys, and reading tests for the REALM Project may be administered to students in a whole group setting (preferred), in small groups, or individually. Please adhere to the following guidelines:

- Ensure all students have testing materials (i.e., correct booklet and a sharpened pencil).
- Ensure that students and pre-labeled booklets are properly paired (i.e., each student receives the booklet that has his or her name on it).
- Provide students with a comfortable testing environment.
- Students are permitted to use mathematics manipulatives during the EMSA if they would ordinarily be permitted to use manipulatives in your classroom.
- No talking or communication between students is permitted during testing.
- Testing administrators should adhere to the guidelines and administration instructions.
- The EMSA, Student Surveys, and reading tests are intended to be read aloud to students by the testing administrator.
- The EMSA administration script indicates that teachers should read the question two times. However, it is permissible for teachers to read the problem more than two times if needed. To ensure uniformity of testing, if the problem is read more than two times, the teacher should read the entire problem to the students and not parts of the problem.

Administering the EMSA, Student Surveys, and Reading Test

It is assumed that the classroom teacher will administer the EMSA, Student Surveys, and reading test; however, other school personnel (such as a paraprofessional or even a substitute teacher) can administer them, providing they follow the assessment protocol as described below. Each packet can be administered on different days or at different times as long as all assessments and surveys are completed during the testing window. Record the testing administration date as well as the start and ending times for both the EMSA test and the Student Surveys on the front page of this administration guide.

The testing conditions should be consistent with the testing conditions for other student assessments administered in the classroom. For example, students should space out the desks or use student "privacy folders" if that is what they would usually do.

Avoid reading problems or answering student questions in a way that may offer clues to the correct answer. Student responses should reflect their current math knowledge. To ensure that the students' test responses are valid, it is important that appropriate procedures are followed when administering the EMSA, Student Surveys, and reading test. These procedures include:

- Administration of the appropriate test level (Kindergarten assessment for Grade K students, etc.)
- Adherence to the guidelines and administration instructions in order to provide a standardized testing protocol across classrooms
- Maintenance of test security

Accommodations

Students with special academic plans (e.g., IEP, 504, ELL) may receive whatever accommodations are specified in their plans.

Testing Time Allocation

Administration of the EMSA, student surveys, and reading test should take no more than 90 minutes combined. These are not timed assessments, and students should be allowed adequate time to answer the questions.

Submitting Materials

Upon conclusion of testing, repack all used and unused EMSA packets (blue paper) and Student Survey/reading test packets (green paper) in their original envelope. Repack any parent permission form packets (yellow paper) in the envelope they came in. Also, please be sure to include the REALM Study Administration Guide (this document) in the envelope. A member of the project team will come collect the testing materials from your school beginning on May 6, 2019. Please have all materials in the front office. You will receive an email prior to pick-up to ensure that all materials are ready in the front office.

If you have questions about this process, contact Charity Bauduin (cbauduin2@lsi.fsu.edu).

Please turn to the next page for the EMSA Administration Script. Instructions for administering the Student Surveys are found on page 17 and for the reading test on page 29.

**Spring 2019 EMSA Administration Instructions
(Blue Booklets)**

This page is intentionally left blank.

EMSA Spring Administration Instructions – Kindergarten

[The boxes contain the script that you will read to students.]

You are about to take a math assessment for a research project at Florida State University. You will need a pencil.

Verify that all students have a pencil.

Before we start the math assessment, I am going to read you a letter from scientists at Florida State University. The letter will explain their research project and will ask if you want to be part of it.

Your teacher is taking part in a research study about how teachers teach math, how students feel about math, and how students learn in math. The people doing the research would like to have copies of your answers on some tests you will take and some questions your teacher will ask you. For some students, we will also observe your class on a few days.

If you agree to be part of this research, you will answer questions about how you feel about math and other things. You will also do some math problems and other tests. This will take about 60-90 minutes on three different days.

You might think some of the questions are too hard. That is ok, we just want you to try your best on them. This study may help teachers learn better ways to help students learn math. We will also ask your school for your scores on some tests you normally take each year.

As a thank you for helping us out you will get a small gift.

Your parents have given permission for you to do this study. But even if your parents gave permission, you can decide to not do the study and that is ok. You do not have to say yes if you do not want the researchers to have this information. Your choice to participate will not affect your grades in school. You can also decide to stop at any time.

You can ask any questions that you have about this study. If you have questions later, you can contact Dr. Colleen Ganley at cganley@fsu.edu with your questions.

If you want to be a part of this research project, say “yes, ” which means that you agree to be in this study and to do the tasks listed above.

If you have a question or aren’t sure if you want to do it, you can raise your hand.

Address any questions. Make sure to receive verbal assent from each student (with a “yes”).

Occasionally a student does not assent, if this happens, you can let them know that the researchers will not look at their answers. Please make a note of any student who do not assent and return the note with all testing materials.

I will now pass out the math assessments. The assessments are already labeled with your names. When you receive the assessment, keep it face up, and do not turn any pages; we will all begin at the same time after I go over the instructions.

Ensure that students and pre-labeled test booklets are properly paired (i.e., each student receives the test booklet that has his or her name written on it). If your students would ordinarily be permitted to use manipulatives and/or scratch paper in this type of situation, ensure that they are available at this time or remind them of your policies for how to access them.

Before we start, please write today’s date, which is _____.

Check to make sure students are writing the correct date.

Turn the page. You should see a pencil in the corner. Look at the first example. It says: What grade are you in? The correct answer choice is K, for Kindergarten. Notice how the bubble below the K has been filled in for you. If you see a question with bubbles below answers, you are going to mark your answer choices the same way, by filling in the bubble below the answer choice you think is correct.

For some problems, you will draw a circle around the answer.

Look at the second example. It says, ‘Draw a circle around the picture of the car.’ The correct answer is circled for you. For some problems, you are going to draw a circle around your answer the same way, by drawing a circle around the answer you think is correct.

For some problems, you will write your answer in a box.

Look at the third example. It says, “Write the number four in the box.” The correct answer is written for you in the box. For some problems, you are going to write your answer in the same way, by writing your answer in a box.

For each problem, I would like you to try hard to figure out which answer is correct. If you are not sure, mark the answer you think is best.

I will read all of the problems to you. Please do not say any answers out loud. You will answer all of the questions by writing on your paper.

You may underline words in the problems if you find that helpful. Also, you can use the blank space on the paper to work out your answers.

Are there any questions?

Address any questions.

I will read each problem two times. If there are no more questions, turn to the page with the bicycle at the top.

Pause; check to ensure all students are on the correct page.

The first problem is, [REDACTED] ?

Fill in the bubble under the answer you think is correct.

I am going to read the problem one more time: [REDACTED] ? Fill in the bubble under the answer you think is correct.

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

Turn to the page with the frog at the top.

Pause; check to ensure all students are on the correct page.

[Redacted]

Fill in the bubble under the answer you think is correct.

I am going to read the problem one more time: [Redacted]

[Redacted] Fill in the bubble under the answer you think is correct.

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

Turn to the page with the car at the top.

Pause; check to ensure all students are on the correct page.

[Redacted] ?

Fill in the bubble under the answer you think is correct.

I am going to read the problem one more time: [Redacted]

[Redacted] ? Fill in the bubble under the answer you think is correct.

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

Turn to the page with the book at the top.

Pause; check to ensure all students are on the correct page.

For this problem you will write your answer in the box.

Write your answer in the box.

I am going to read the problem one more time:

Write your answer in the box.

the box.

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

Turn to the page with the soccer ball at the top.

Pause; check to ensure all students are on the correct page.

For the next four problems I am going to read an equation aloud to you. If the equation is correct, circle the word yes. If the equation is not correct, circle the word no.

The first equation is . If the equation is correct, circle the word yes. If the equation is not correct, circle the word no.

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

The last equation on this page is . Again, . If the equation is correct, circle the word yes. If the equation is not correct, circle the word no.

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

Turn to the page with the apple at the top.

Pause; check to ensure all students are on the correct page.

The first equation is [REDACTED]. Again, [REDACTED]. If the equation is correct, circle the word yes. If the equation is not correct, circle the word no.

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

The last equation on this page is [REDACTED]. Again, [REDACTED]. If the equation is correct, circle the word yes. If the equation is not correct, circle the word no.

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

Turn to the page with the dog at the top.

Pause; check to ensure all students are on the correct page.

[REDACTED].
I am going to read the problem one more time. [REDACTED]
[REDACTED].

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

Turn to the page with the scissors at the top.

Pause; check to ensure all students are on the correct page.

[REDACTED].

I am going to read the problem one more time. [REDACTED]

[REDACTED].

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

Turn to the page with the smiley face at the top.

Pause; check to ensure all students are on the correct page.

There are two problems on this page. The first problem is [REDACTED]. Fill in the bubble below the number you think should go in the box.

I am going to read the problem one more time. [REDACTED] **. Fill in the bubble below the number you think should go in the box.**

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

The next problem is [REDACTED]. Fill in the bubble below the number you think should go in the box.

I am going to read the problem one more time. [REDACTED] **. Fill in the bubble below the number you think should go in the box.**

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

Turn to the page with the balloon at the top.

Pause; check to ensure all students are on the correct page.

[Redacted]. Fill in the bubble below the number you think should go in the box.

I am going to read the problem one more time. **[Redacted]**. Fill in the bubble below the number you think should go in the box.

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

Turn to the page with the fish at the top.

Pause; check to ensure all students are on the correct page.

[Redacted] ? Fill in the bubble under the answer you think is correct.

I am going to read the problem one more time. **[Redacted]** ? Fill in the bubble under the answer you think is correct.

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

Turn to the page with the zebra at the top.

Pause; check to ensure all students are on the correct page.

[Redacted] ? Fill in the bubble under the answer you think is correct.

I am going to read the problem one more time. **[Redacted]** ? Fill in the bubble answer you think is correct.

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

Turn to the page with the chair at the top.

Pause; check to ensure all students are on the correct page.

There are two problems on this page. The first problem is [REDACTED]. Write the number that goes in the box to make this equation correct.

I am going to read the problem one more time. [REDACTED]. Write the number that goes in the box to make this equation correct.

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

The last problem is [REDACTED]. Write the number that goes in the box to make this equation correct.

I am going to read the problem one more time. [REDACTED]. Write the number that goes in the box to make this equation correct.

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

We are all finished with this assessment. I am going to collect all the booklets.

Collect all testing materials.

Appendix F. Grade 1 Administration Guide

The form in this appendix is identical to the form used in spring 2019. As a result, there are no headers or footers present in this section of the report. The EMSA test was administered with two other assessments and those assessments are reflected in the contents of the administration guide.

Administration Instructions for the REALM Project

Grade 1 Spring Assessments and Surveys

2018–2019

Teacher Name: _____

EMSA Tests (Blue Booklet)

Administration Date: _____

Start Time: _____ **End Time:** _____

Student Surveys (Green Booklet)

Administration Date: _____

Start Time: _____ **End Time:** _____

This page is intentionally left blank.

Table of Contents

Administration Guidelines for Assessments and Student Surveys	1
Overview	1
Materials	1
EMSA Test, Surveys, and Reading Booklets.....	1
Students to be Tested	2
Preparing for Testing	2
Administering the EMSA, Student Surveys, and Reading Test	4
Accommodations.....	4
Testing Time Allocation	4
Submitting Materials	4
Spring 2019 EMSA Administration Instructions.....	5

This page is intentionally left blank.

Administration Guidelines for Assessments and Student Surveys

Overview

Thank you for your participation in the REALM Project. This document will provide you with instructions to follow for the purpose of assessing your mathematics students. Within this document are administration guidelines and instructions for assessments and student surveys. There is also important information about the testing window, as well as returning all materials to the REALM Project staff.

The assessments and student surveys are designed to be administered in a written format with the whole class, but you may choose to administer individually or in small groups as you see fit. Please administer the Spring 2019 Elementary Mathematics Student Assessment (EMSA), student surveys, and reading test during the allotted testing window of April 15, 2019 to May 3, 2019. If you cannot administer the assessments during that window, please notify Charity Bauduin (cbauduin2@lsi.fsu.edu) and plan to administer them as soon as possible. You may choose to administer the EMSA test, student surveys, and reading test on separate days. If you choose to administer all assessments and surveys in the same day, you may consider providing students with a brief stretch break in between each assessment. We do ask that you administer the EMSA first, followed by the student surveys, and conclude with the reading test.

Scripts for the teacher to use during administration begin on page 7 of this guide. Please follow the scripts as closely as possible when you or your surrogate administer the assessments and surveys. assessments, surveys, and administration guides will be picked up as described in the Submitting the Materials section on page 4.

We ask that students use pencils to complete these assessments.

Materials

The following materials are required for testing:

- Administration Instructions for the Grade 1 Spring Assessments and Surveys (this document)
- An EMSA booklet for each student (blue packet, one per student, provided)
- A Student Surveys booklet for each student (green packet, one per student, provided)
- A Reading test booklet for each student (one per student, provided)
- Student Assent (one each for students who did not provide assent in the fall, provided)
- A FSU REALM Project pencil for each student (one per student, provided)
- At least one sharpened pencil for each student

EMSA Test, Surveys, and Reading Booklets

Students should mark their answers directly in the booklets. Should you need additional testing materials, please contact Charity Bauduin (cbauduin2@lsi.fsu.edu). Remember that these materials are to remain at the school site until the testing window has ended. The materials should be stored in a secure, access-restricted location at all times.

Students to be Tested

We ask that you administer the math assessment (EMSA), student surveys, and reading test to students for whom you are the teacher of record. If you teach multiple groups of students mathematics, please contact Charity Bauduin (cbauduin2@lsi.fsu.edu) for additional testing materials. We have provided enough materials so that all students may complete the assessments and surveys. However, we will only enter and analyze data for those children who meet the requirements for consent.

Preparing for Testing

The first page of each EMSA booklet (blue packet) has the following box for student information:

Date:	
District:	School
Teacher:	
Student:	Grade

We have attached labels for those students who were included on the class rosters provided to us. If you have students with pre-labeled booklets that have moved, please cross out their names. Please use these booklets for any student who does not have pre-labeled booklets. There are additional booklets containing no labels that can also be used for students who do not have pre-labeled booklets. For those students who do not have pre-labeled EMSA tests, we ask that the classroom teacher, or an adult surrogate, enter this information (date, district name, school name, teacher name, student full name as it appears on official records, and student grade level) on each test booklet for each student to be tested prior to the testing session. (Please do not leave it for students to enter this information.)

The first page of each Student Surveys booklet (green packet) has the following box for student information:

Student’s Name:	Date:
Teacher:	

The first page of each reading test booklet (purple packet) has the following lines for student information:

Student's Name: _____

Date: _____

Prior to the testing session, the classroom teacher, or an adult surrogate, must enter this information (student's name full name as it appears on official records, date, and teacher name) on each survey and reading booklet for each student to be tested that does not have a pre-labeled booklet. (Please do not leave it for students to enter this information.)

The Spring EMSA, Student Surveys, and reading tests for the REALM Project may be administered to students in a whole group setting (preferred), in small groups, or individually. Please adhere to the following guidelines:

- Ensure all students have testing materials (i.e., correct booklet and a sharpened pencil).
- Ensure that students and pre-labeled booklets are properly paired (i.e., each student receives the booklet that has his or her name on it).
- Provide students with a comfortable testing environment.
- Students are permitted to use mathematics manipulatives during the EMSA if they would ordinarily be permitted to use manipulatives in your classroom.
- No talking or communication between students is permitted during testing.
- Testing administrators should adhere to the guidelines and administration instructions.
- The EMSA, Student Surveys, and reading tests are intended to be read aloud to students by the testing administrator.
- The EMSA administration script indicates that teachers should read the question two times. However, it is permissible for teachers to read the problem more than two times if needed. To ensure uniformity of testing, if the problem is read more than two times, the teacher should read the entire problem to the students and not parts of the problem.

Administering the EMSA, Student Surveys, and Reading Test

It is assumed that the classroom teacher will administer the EMSA, Student Surveys, and reading test; however, other school personnel (such as a paraprofessional or even a substitute teacher) can administer them, providing they follow the assessment protocol as described below. Each packet can be administered on different days or at different times as long as all assessments and surveys are completed during the testing window. Record the testing administration date as well as the start and ending times for both the EMSA test and the Student Surveys on the front page of this administration guide.

The testing conditions should be consistent with the testing conditions for other student assessments administered in the classroom. For example, students should space out the desks or use student “privacy folders” if that is what they would usually do.

Avoid reading problems or answering student questions in a way that may offer clues to the correct answer. Student responses should reflect their current math knowledge. To ensure that the students’ test responses are valid, it is important that appropriate procedures are followed when administering the EMSA, Student Surveys, and reading test. These procedures include:

- Administration of the appropriate test level (Grade 1 assessment for Grade 1 students, etc.)
- Adherence to the guidelines and administration instructions in order to provide a standardized testing protocol across classrooms
- Maintenance of test security

Accommodations

Students with special academic plans (e.g., IEP, 504, ELL) may receive whatever accommodations are specified in their plans.

Testing Time Allocation

Administration of the EMSA, student surveys, and reading test should take no more than 90 minutes combined. These are not timed assessments, and students should be allowed adequate time to answer the questions.

Submitting Materials

Upon conclusion of testing, repack all used and unused EMSA packets (blue paper), Student Survey (green paper), and reading assessment packets in their original envelope. Repack any parent permission form packets (yellow paper) and student assent forms in the envelope they came in. Also, please be sure to include the REALM Study Administration Guide (this document) in the envelope. A member of the project team will come collect the testing materials from your school beginning on May 6, 2019. Please have all materials in the front office. You will receive an email prior to pick-up to ensure that all materials are ready in the front office.

If you have questions about this process, contact Charity Bauduin (cbauduin2@lsi.fsu.edu).

Please turn to the next page for the EMSA Administration Script. Instructions for administering the Student Surveys are found on page 17 and for the reading test on page 33.

**Spring 2019 EMSA Administration Instructions
(Blue Booklets)**

This page is intentionally left blank.

EMSA Spring Administration Instructions – Grade 1

[The boxes contain the script that you will read to students.]

You are about to take a math assessment for a research project at Florida State University. You will need a pencil.

Verify that all students have a pencil. Pass out the student assent forms to those students who have forms labeled for them.

Before we start the math assessment, I am going to read you a letter from scientists at Florida State University. The letter will explain their research project and will ask if you want to be part of it. Some of you have already signed this letter in the fall and will not need to sign it again.

Your teacher is participating in a research study about how teachers teach math, how students feel about math, and how students learn in math. The people doing the research would like to have copies of your answers on some tests that you will take and some questions your teacher will ask you. For some students, we will also observe your classroom on a few days.

If you agree to be part of this research, you will answer questions about how you feel about math and other subjects and also do some math problems and other tests. This will take a total of about 60-90 minutes at three different times. You might think some of the questions are too hard, but that is ok, we just want you to try your best on them. This study may help teachers learn better ways to help students with their math skills. We will also ask your school for your scores on some tests you normally take each year.

As a thank you for helping us out, you will get a small gift.

Your parents have given their permission for you to do this study. But even if your parents gave permission, you can still decide to not do the study and that is ok. You do not have to sign this paper if you do not want the researchers to have this information. Your choice to participate in the research study will not affect your grades in school.

You can ask any questions that you have about this study. If you have questions later, you can contact Dr. Colleen Ganley at cganley@fsu.edu with your questions.

Signing your name at the bottom means that you agree to be in this study and to do the tasks listed above.

If you want to be a part of this research project, write your name on the line and then write today's date which is _____.

If you have a question, you can raise your hand.

Address any questions.

Verify that all students with an assent form have written their first and last name (if possible) on the assent form. You may help students by writing their name if they give you verbal assent. Occasionally a student does not assent, if this happens, you can let them know that the researchers will not look at their answers. Please collect all assent forms.

I will now pass out the math assessments. The assessments are already labeled with your names. When you receive the assessment, keep it face up, and do not turn any pages; we will all begin at the same time after I go over the instructions. Some of these questions may be hard, but don't worry and just try your best.

Ensure that students and pre-labeled test booklets are properly paired (i.e., each student receives the test booklet that has his or her name written on it). If your students would ordinarily be permitted to use manipulatives and/or scratch paper in this type of situation, ensure that they are available at this time or remind them of your policies for how to access them.

Before we start, please write today's date, which is _____.

Check to make sure students are writing the correct date.

Turn the page. You should see a pencil in the corner. Look at the first example. It says: What grade are you in? The correct answer choice is 1, for first grade. Notice how the bubble below the 1 has been filled in for you. If you see a question with bubbles below answers, you are going to mark your answer choices the same way, by filling in the bubble below the answer choice you think is correct.

For some problems, you will draw a circle around the answer you think is correct.

Look at the second example. It says, 'Draw a circle around the picture of the car.' The correct answer is circled for you. For some problems, you are going to draw a circle around your answer the same way, by drawing a circle around the answer you think is correct.

For some problems, you will write your answer in a box.

Look at the third example. It says, "Write the number four in the box." The correct answer is written for you in the box. For some problems, you are going to write your answer in the same way, by writing your answer in a box.

For each problem, I would like you to try hard to figure out which answer is correct. If you are not sure, mark the answer you think is best.

I will read all of the problems to you. Please do not say any answers out loud. You will answer all of the questions by writing on your paper.

You may underline words in the problems if you find that helpful. Also, you can use the blank space on the paper to work out your answers.

Are there any questions?

Address any questions.

I will read each problem two times. If there are no more questions, turn to the page with the frog at the top.

Pause; check to ensure all students are on the correct page.

The first problem is,

?

Fill in the bubble under the answer you think is correct.

I am going to read the problem one more time:

?

Fill in the bubble under the answer you think is correct.

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

Turn to the page with the car at the top.

Pause; check to ensure all students are on the correct page.

For this problem you will write your answer in the box.

? Write your answer in the box.

I am going to read the problem one more time:

? Write your answer in the box.

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

Turn to the page with the bicycle at the top.

Pause; check to ensure all students are on the correct page.

? Write your answer in the box.

I am going to read the problem one more time:

? Write your answer in the box.

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

Turn to the page with the soccer ball at the top.

Pause; check to ensure all students are on the correct page.

? Write your answer in the box.

I am going to read the problem one more time:

Write your answer in the box.

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

Turn to the page with the book at the top.

Pause; check to ensure all students are on the correct page.

For the next five problems I am going to read an equation aloud to you. If the equation is correct, circle the word yes. If the equation is not correct, circle the word no.

The first equation is . Again, . If the equation is correct, circle the word yes. If the equation is not correct, circle the word no.

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

The last equation on this page is . Again, . If the equation is correct, circle the word yes. If the equation is not correct, circle the word no.

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

Turn to the page with the bus at the top.

Pause; check to ensure all students are on the correct page.

The first equation on this page is [REDACTED]. Again, [REDACTED]. If the equation is correct, circle the word yes. If the equation is not correct, circle the word no.

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

The last equation on this page is [REDACTED]. Again, [REDACTED]. If the equation is correct, circle the word yes. If the equation is not correct, circle the word no.

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

Turn to the page with the chair at the top.

Pause; check to ensure all students are on the correct page.

The equation on this page is [REDACTED]. Again, [REDACTED]. If the equation is correct, circle the word yes. If the equation is not correct, circle the word no.

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

Turn to the page with the dog at the top.

Pause; check to ensure all students are on the correct page.

[REDACTED].

I am going to read the problem one more time. [REDACTED].

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

Turn to the page with the scissors at the top.

Pause; check to ensure all students are on the correct page.

I am going to read the problem one more time.

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

Turn to the page with the apple at the top.

Pause; check to ensure all students are on the correct page.

I am going to read the problem one more time.

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

Turn to the page with the fish at the top.

Pause; check to ensure all students are on the correct page.

[REDACTED]

Fill in the bubble under the answer you think is correct.

I am going to read the problem one more time: [REDACTED]

[REDACTED] Fill in the bubble under the answer you think is correct.

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

Turn to the page with the zebra at the top.

Pause; check to ensure all students are on the correct page.

There are two problems on this page. The first problem is [REDACTED]. Fill in the bubble below the number you think should go in the box.

I am going to read the problem one more time. [REDACTED]. Fill in the bubble below the number you think should go in the box.

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

The next problem is [REDACTED]. Fill in the bubble below the number you think should go in the box.

I am going to read the problem one more time. [REDACTED]. Fill in the bubble below the number you think should go in the box.

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

Turn to the page with the smiley face at the top.

Pause; check to ensure all students are on the correct page.

There are two problems on this page. The first problem is [REDACTED].

Fill in the bubble under the answer you think is correct.

I am going to read the problem one more time. [REDACTED] ? Fill in the bubble under the answer you think is correct.

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

The last problem on this page is [REDACTED] ? Write your answer in the box.

I am going to read the problem one more time. [REDACTED] ? Write your answer in the box.

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

Turn to the page with the backpack at the top.

Pause; check to ensure all students are on the correct page.

Now you are going to work on some addition and subtraction problems on your own. There are three problems on this page and two problems on the next page that you will solve at your own pace. Your job is to find out what number goes in the box to make the equation correct. Write your answer for each problem in the box.

Remember to look closely at the symbol to decide if it is an addition or subtraction problem. When I say “begin” you can start answering the questions. When you get to the end of the first page, continue on to the next page until you reach the stop sign at the bottom of the last page. Are there any questions?

Address any questions.

You may begin.

Circulate as students work on the problems. Provide students with ample time to complete the problems. Once you see that all students have completed the problems, please end the assessment.

We are all finished with this assessment. I am going to collect all the booklets.

Collect all testing materials.

Appendix G. Grade 2 Administration Guide

The form in this appendix is identical to the form used in spring 2019. As a result, there are no headers or footers present in this section of the report. The EMSA test was administered with two other assessments and those assessments are reflected in the contents of the administration guide.

Administration Instructions for the REALM Project Grade 2 Spring Assessments and Surveys

2018–2019

Teacher Name: _____

EMSA Tests (Blue Booklet)

Administration Date: _____

Start Time: _____ **End Time:** _____

Student Surveys (Green Booklet)

Administration Date: _____

Start Time: _____ **End Time:** _____

This page is intentionally left blank.

Table of Contents

Administration Guidelines for Assessments and Student Surveys	1
Overview	1
Materials	1
EMSA Test, Surveys, and Reading Booklets.....	1
Students to be Tested	2
Preparing for Testing	2
Administering the EMSA, Student Surveys, and Reading Test	4
Accommodations.....	4
Testing Time Allocation	4
Submitting Materials	4
Spring 2019 EMSA Administration Instructions.....	5

This page is intentionally left blank.

Administration Guidelines for Assessments and Student Surveys

Overview

Thank you for your participation in the REALM Project. This document will provide you with instructions to follow for the purpose of assessing your mathematics students. Within this document are administration guidelines and instructions for assessments and student surveys. There is also important information about the testing window, as well as returning all materials to the REALM Project staff.

The assessments and student surveys are designed to be administered in a written format with the whole class, but you may choose to administer individually or in small groups as you see fit. Please administer the Spring 2019 Elementary Mathematics Student Assessment (EMSA), student surveys, and reading test during the allotted testing window of April 15, 2019 to May 3, 2019. If you cannot administer the assessments during that window, please notify Charity Bauduin (cbauduin2@lsi.fsu.edu) and plan to administer them as soon as possible. You may choose to administer the EMSA test, student surveys, and reading test on separate days. If you choose to administer all assessments and surveys in the same day, you may consider providing students with a brief stretch break in between each assessment. We do ask that you administer the EMSA first, followed by the student surveys, and conclude with the reading test.

Scripts for the teacher to use during administration begin on page 7 of this guide. Please follow the scripts as closely as possible when you or your surrogate administer the assessments and surveys. The assessments, surveys, and administration guides will be picked up as described in the Submitting the Materials section on page 4.

We ask that students use pencils to complete these assessments.

Materials

The following materials are required for testing:

- Administration Instructions for the Grade 2 Spring Assessments and Surveys (this document)
- An EMSA booklet for each student (blue packet, one per student, provided)
- A Student Surveys booklet for each student (green packet, one per student, provided)
- A Reading test booklet for each student (one per student, provided)
- Student Assent (one per student, provided)
- A FSU REALM Project pencil for each student (one per student, provided)
- At least one sharpened pencil for each student

EMSA Test, Surveys, and Reading Booklets

Students should mark their answers directly in the booklets. Should you need additional testing materials, please contact Charity Bauduin (cbauduin2@lsi.fsu.edu). Remember that these materials are to remain at the school site until the testing window has ended. The materials should be stored in a secure, access-restricted location at all times.

Students to be Tested

We ask that you administer the math assessment (EMSA), student surveys, and reading test to students for whom you are the teacher of record. If you teach multiple groups of students mathematics, please contact Charity Bauduin (cbauduin2@lsi.fsu.edu) for additional testing materials. We have provided enough materials so that all students may complete the assessments and surveys. However, we will only enter and analyze data for those children who meet the requirements for consent.

Preparing for Testing

The first page of each EMSA booklet (blue packet) has the following box for student information:

Date:	
District:	School
Teacher:	
Student:	Grade

We have attached labels for those students who were included on the class rosters provided to us. If you have students with pre-labeled booklets that have moved, please cross out their names. Please use these booklets for any student who does not have pre-labeled booklets. There are additional booklets containing no labels that can also be used for students who do not have pre-labeled booklets. For those students who do not have pre-labeled EMSA tests, we ask that the classroom teacher, or an adult surrogate, enter this information (date, district name, school name, teacher name, student full name as it appears on official records, and student grade level) on each test booklet for each student to be tested prior to the testing session. (Please do not leave it for students to enter this information.)

The first page of each Student Surveys booklet (green packet) has the following box for student information:

Student's Name:	Date:
Teacher:	

The first page of each reading test booklet (red packet) has the following lines for student information:

Student's Name: _____

Date: _____

Prior to the testing session, the classroom teacher, or an adult surrogate, must enter this information (student's name full name as it appears on official records, date, and teacher name) on each survey and reading booklet for each student to be tested who does not have a pre-labeled booklet. (Please do not leave it for students to enter this information.)

The Spring EMSA, Student Surveys, and reading tests for the REALM Project may be administered to students in a whole group setting (preferred), in small groups, or individually. Please adhere to the following guidelines:

- Ensure all students have testing materials (i.e., correct booklet and a sharpened pencil).
- Ensure that students and pre-labeled booklets are properly paired (i.e., each student receives the booklet that has his or her name on it).
- Provide students with a comfortable testing environment.
- Students are permitted to use mathematics manipulatives during the EMSA if they would ordinarily be permitted to use manipulatives in your classroom.
- No talking or communication between students is permitted during testing.
- Testing administrators should adhere to the guidelines and administration instructions.
- The EMSA, Student Surveys, and reading tests are intended to be read aloud to students by the testing administrator.
- The EMSA administration script indicates that teachers should read the question two times. However, it is permissible for teachers to read the problem more than two times if needed. To ensure uniformity of testing, if the problem is read more than two times, the teacher should read the entire problem to the students and not parts of the problem.

Administering the EMSA, Student Surveys, and Reading Test

It is assumed that the classroom teacher will administer the EMSA, Student Surveys, and reading test; however, other school personnel (such as a paraprofessional or even a substitute teacher) can administer them, providing they follow the assessment protocol as described below. Each packet can be administered on different days or at different times as long as all assessments and surveys are completed during the testing window. Record the testing administration date as well as the start and ending times for both the EMSA test and the Student Surveys on the front page of this administration guide.

The testing conditions should be consistent with the testing conditions for other student assessments administered in the classroom. For example, students should space out the desks or use student “privacy folders” if that is what they would usually do.

Avoid reading problems or answering student questions in a way that may offer clues to the correct answer. Student responses should reflect their current math knowledge. To ensure that the students’ test responses are valid, it is important that appropriate procedures are followed when administering the EMSA, Student Surveys, and reading test. These procedures include:

- Administration of the appropriate test level (Grade 2 assessment for Grade 2 students, etc.)
- Adherence to the guidelines and administration instructions in order to provide a standardized testing protocol across classrooms
- Maintenance of test security

Accommodations

Students with special academic plans (e.g., IEP, 504, ELL) may receive whatever accommodations are specified in their plans.

Testing Time Allocation

Administration of the EMSA, student surveys, and reading test should take no more than 90 minutes combined. These are not timed assessments, and students should be allowed adequate time to answer the questions.

Submitting Materials

Upon conclusion of testing, repack all used and unused EMSA packets (blue paper), Student Survey (green paper), and reading assessment packets in their original envelope. Repack parent permission form packets (yellow paper) and student assents in the envelope they came in. Also, please be sure to include the REALM Study Administration Guide (this document) in the envelope. A member of the project team will come collect the testing materials from your school on May 6, 2019. Please have all materials in the front office. You will receive an email prior to pick-up to ensure that all materials are ready in the front office.

If you have questions about this process, contact Charity Bauduin (cbauduin2@lsi.fsu.edu).

Please turn to the next page for the EMSA Administration Script. Instructions for administering the Student Surveys are found on page 19 and for the reading test on page 33.

**Spring 2019 EMSA Administration Instructions
(Blue Booklets)**

This page is intentionally left blank.

EMSA Spring Administration Instructions – Grade 2

[The boxes contain the script that you will read to students.]

You are about to take a math assessment for a research project at Florida State University. You will need a pencil.

Verify that all students have a pencil. Pass out the student assent forms to those students who have forms labeled for them. There are additional blank forms for any new students in your class.

Before we start the math assessment, I am going to read you a letter from scientists at Florida State University. The letter will explain their research project and will ask if you want to be part of it. Some of you have already signed this letter in the fall and will not need to sign it again.

Your teacher is participating in a research study about how teachers teach math, how students feel about math, and how students learn in math. The people doing the research would like to have copies of your answers on some tests that you will take and some questions your teacher will ask you. For some students, we will also observe your classroom on a few days.

If you agree to be part of this research, you will answer questions about how you feel about math and other subjects and also do some math problems and other tests. This will take a total of about 60-90 minutes at three different times. You might think some of the questions are too hard, but that is ok, we just want you to try your best on them. This study may help teachers learn better ways to help students with their math skills. We will also ask your school for your scores on some tests you normally take each year.

As a thank you for helping us out, you will get a small gift.

Your parents have given their permission for you to do this study. But even if your parents gave permission, you can still decide to not do the study and that is ok. You do not have to sign this paper if you do not want the researchers to have this information. Your choice to participate in the research study will not affect your grades in school.

You can ask any questions that you have about this study. If you have questions later, you can contact Dr. Colleen Ganley at cganley@fsu.edu with your questions.

Signing your name at the bottom means that you agree to be in this study and to do the tasks listed above.

If you want to be a part of this research project, write your name on the line and then write today's date which is _____

If you have a question, you can raise your hand.

Address any questions.

Verify that all students with an assent form have written their first and last name (if possible) on the assent form. You may help students by writing their name if they give you verbal assent. Occasionally a student does not assent, if this happens, you can let them know that the researchers will not look at their answers. Please collect all assent forms.

I will now pass out the math assessments. The assessments are already labeled with your names. When you receive the assessment, keep it face up, and do not turn any pages; we will all begin at the same time after I go over the instructions. Some of these questions may be hard, but don't worry and just try your best.

Ensure that students and pre-labeled test booklets are properly paired (i.e., each student receives the test booklet that has his or her name written on it). If your students would ordinarily be permitted to use manipulatives and/or scratch paper in this type of situation, ensure that they are available at this time or remind them of your policies for how to access them.

Before we start, please write today's date, which is _____.

Check to make sure students are writing the correct date.

Turn the page. You should see a pencil in the corner. Look at the first example. It says: What grade are you in? The correct answer choice is 2, for second grade. Notice how the bubble below the 2 has been filled in for you. If you see a question with bubbles below answers, you are going to mark your answer choices the same way, by filling in the bubble below the answer choice you think is correct.

For some problems, you will draw a circle around the answer you think is correct.

Look at the second example. It says, 'Draw a circle around the picture of the car.' The correct answer is circled for you. For some problems, you are going to draw a circle around your answer the same way, by drawing a circle around the answer you think is correct.

For some problems, you will write your answer in a box.

Look at the third example. It says, "Write the number four in the box." The correct answer is written for you in the box. For some problems, you are going to write your answer in the same way, by writing your answer in a box.

For each problem, I would like you to try hard to figure out which answer is correct. If you are not sure, mark the answer you think is best.

I will read all of the problems to you. Please do not say any answers out loud. You will answer all of the questions by writing on your paper.

You may underline words in the problems if you find that helpful. Also, you can use the blank space on the paper to work out your answers.

Are there any questions?

Address any questions.

I will read each problem two times. Turn to the page with the frog at the top.

Pause; check to ensure all students are on the correct page.

The first problem is,

[REDACTED]

I am going to read the problem one more time:

[REDACTED]

? Write your answer in the box.

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

Turn to the page with the dog at the top.

Pause; check to ensure all students are on the correct page.

? Write your answer in the box.

I am going to read the problem one more time:

? Write your answer in the box.

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

Turn to the page with the soccer ball at the top.

Pause; check to ensure all students are on the correct page.

Write your answer in the box.

I am going to read the problem one more time:

? Write your answer in the box.

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

Turn to the page with the car at the top.

Pause; check to ensure all students are on the correct page.

? Write your answer in the box.

I am going to read the problem one more time:

answer in the box.

Write your

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

Turn to the page with the zebra at the top.

Pause; check to ensure all students are on the correct page.

? Write your answer in the box.

I am going to read the problem one more time:

the box.

? Write your answer in

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

Turn to the page with the balloon at the top.

Pause; check to ensure all students are on the correct page.

For the next five problems I am going to read an equation aloud to you. If the equation is correct, circle the word yes. If the equation is not correct, circle the word no.

The first equation is [REDACTED]. Again, [REDACTED]. If the equation is correct, circle the word yes. If the equation is not correct, circle the word no.

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

The last equation on this page is [REDACTED]. Again, [REDACTED]. If the equation is correct, circle the word yes. If the equation is not correct, circle the word no.

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

Turn to the page with the chair at the top.

Pause; check to ensure all students are on the correct page.

The first equation on this page is [REDACTED]. Again, [REDACTED]. If the equation is correct, circle the word yes. If the equation is not correct, circle the word no.

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

The last equation on this page is [REDACTED]. Again, [REDACTED]. If the equation is correct, circle the word yes. If the equation is not correct, circle the word no.

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

Turn to the page with the bicycle at the top.

Pause; check to ensure all students are on the correct page.

The equation on this page is [REDACTED]. Again, [REDACTED]. If the equation is correct, circle the word yes. If the equation is not correct, circle the word no.

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

Turn to the page with the backpack at the top.

Pause; check to ensure all students are on the correct page.

[REDACTED].
I am going to read the problem one more time. [REDACTED].

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

Turn to the page with the smiley face at the top.

Pause; check to ensure all students are on the correct page.

[Redacted]

I am going to read the problem one more time. [Redacted]
[Redacted]

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

Turn to the page with the apple at the top.

Pause; check to ensure all students are on the correct page.

[Redacted]

I am going to read the problem one more time. [Redacted]
[Redacted]

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

Turn to the page with the bus at the top.

Pause; check to ensure all students are on the correct page.

[Redacted]

Fill in the bubble under the answer you think is correct.

I am going to read the problem one more time: [Redacted]
[Redacted] Fill in the bubble under the answer you think is correct.

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

Turn to the page with the scissors at the top.

Pause; check to ensure all students are on the correct page.

There are two problems on this page. The first problem is [REDACTED]. Fill in the bubble below the number you think should go in the box.

I am going to read the problem one more time. [REDACTED]. Fill in the bubble below the number you think should go in the box.

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

The next problem is [REDACTED].

Fill in the bubble below the number you think should go in the box.

I am going to read the problem one more time. [REDACTED]. Fill in the bubble below the number you think should go in the box.

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

Turn to the page with the cat at the top.

Pause; check to ensure all students are on the correct page.

There are two problems on this page. The first problem is ? Write your answer in the box.

I am going to read the problem one more time. ? Write your answer in the box.

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

The last problem on this page is . Write your answer in the box.

I am going to read the problem one more time. ? Write your answer in the box.

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

Turn to the page with the fish at the top.

Pause; check to ensure all students are on the correct page.

Now you are going to work on some problems on your own. The next three pages have some addition and subtraction problems that you will solve at your own pace. Your job is to find out what number goes in the box to make the equation correct. Write your answer for each problem in the box.

Remember to look closely at the symbol to decide if it is an addition or subtraction problem. When I say “begin” you can start answering the questions. When you get to the end of the first page, continue on to the next page until you reach the stop sign at the bottom of the last page. Are there any questions?

Address any questions.

You may begin.

Circulate as students work on the problems. Provide students with ample time to complete the problems. Once you see that all students have completed the problems, please end the assessment.

Ok, we are all finished with this assessment. I am going to collect all the booklets.

Collect all testing materials.

Appendix H. Grade 3 Administration Guide

The form in this appendix is identical to the form used in spring 2019. As a result, there are no headers or footers present in this section of the report. The EMSA test was administered with two other assessments and those assessments are reflected in the contents of the administration guide.

Administration Instructions for the REALM Project

Grade 3 Spring Assessments and Surveys

2018–2019

Teacher Name: _____

EMSA Tests (Blue Booklet)

Administration Date: _____

Start Time: _____ **End Time:** _____

Student Surveys (Green Booklet)

Administration Date: _____

Start Time: _____ **End Time:** _____

This page is intentionally left blank.

Table of Contents

Administration Guidelines for Assessments and Student Surveys.....	1
Overview.....	1
Materials.....	1
EMSA Test, Surveys, and Reading Booklets	1
Students to be Tested.....	2
Preparing for Testing	2
Administering the EMSA, Student Surveys, and Reading Test.....	4
Accommodations	4
Testing Time Allocation	4
Submitting Materials	4
Spring 2019 EMSA Administration Instructions	5

This page is intentionally left blank.

Administration Guidelines for Assessments and Student Surveys

Overview

Thank you for your participation in the REALM Project. This document will provide you with instructions to follow for the purpose of assessing your mathematics students. Within this document are administration guidelines and instructions for assessments and student surveys. There is also important information about the testing window, as well as returning all materials to the REALM Project staff.

The assessments and student surveys are designed to be administered in a written format with the whole class, but you may choose to administer individually or in small groups as you see fit. Please administer the Spring 2018 Elementary Mathematics Student Assessment (EMSA), student surveys, and reading test during the allotted testing window of April 15, 2019 to May 3, 2019. If you cannot administer the assessments during that window, please notify Charity Bauduin (cbauduin2@lsi.fsu.edu) and plan to administer them as soon as possible. You may choose to administer the EMSA test, student surveys, and reading test on separate days. If you choose to administer all assessments and surveys in the same day, you may consider providing students with a brief stretch break in between each assessment. We do ask that you administer the EMSA first, followed by the student surveys, and conclude with the reading test.

Scripts for the teacher to use during administration begin on page 5 of this guide. Please follow the scripts as closely as possible when you or your surrogate administer the assessments and surveys. The assessments, surveys, and administration guides will be picked up as described in the Submitting the Materials section on page 4.

We ask that students use pencils to complete these assessments.

Materials

The following materials are required for testing:

- Administration Instructions for the Grade 3 Spring Assessments and Surveys (this document)
- An EMSA booklet for each student (blue packet, one per student, provided)
- A Student Surveys booklet for each student (green packet, one per student, provided)
- A Reading test booklet for each student (one per student, provided)
- Student Assent (one per student, provided)
- A FSU REALM Project pencil for each student (one per student, provided)
- At least one sharpened pencil for each student

EMSA Test, Surveys, and Reading Booklets

Students should mark their answers directly in the booklets. Should you need additional testing materials, please contact Charity Bauduin (cbauduin2@lsi.fsu.edu). Remember that these materials are to remain at the school site until the testing window has ended. The materials should be stored in a secure, access-restricted location at all times.

Students to be Tested

We ask that you administer the math assessment (EMSA), student surveys, and reading test to students for whom you are the teacher of record. If you teach multiple groups of students mathematics, please contact Charity Bauduin (cbauduin2@lsi.fsu.edu) for additional testing materials. We have provided enough materials so that all students may complete the assessments and surveys. However, we will only enter and analyze data for those children who meet the requirements for consent.

Preparing for Testing

The first page of each EMSA booklet (blue packet) has the following box for student information:

Date:	
District:	School
Teacher:	
Student:	Grade

We have attached labels for those students who were included on the class rosters provided to us. If you have students with pre-labeled booklets that have moved, please cross out their names. Please use these booklets for any student who does not have pre-labeled booklets. There are additional booklets containing no labels that can also be used for students who do not have pre-labeled booklets. For those students who do not have pre-labeled EMSA tests, we ask that the classroom teacher, or an adult surrogate, enter this information (date, district name, school name, teacher name, student full name as it appears on official records, and student grade level) on each test booklet for each student to be tested prior to the testing session. (Please do not leave it for students to enter this information.)

The first page of each Student Surveys booklet (green packet) has the following box for student information:

Student’s Name:	Date:
Teacher:	

The first page of each reading test booklet (green packet) has the following lines for student information:

Student's Name: _____

Date: _____

Prior to the testing session, the classroom teacher, or an adult surrogate, must enter this information (student's name full name as it appears on official records, date, and teacher name) on each survey and reading booklet for each student to be tested who does not have a pre-labeled booklet. (Please do not leave it for students to enter this information.)

The Spring EMSA, Student Surveys, and reading tests for the REALM Project may be administered to students in a whole group setting (preferred), in small groups, or individually. Please adhere to the following guidelines:

- Ensure all students have testing materials (i.e., correct booklet and a sharpened pencil).
- Ensure that students and pre-labeled booklets are properly paired (i.e., each student receives the booklet that has his or her name on it).
- Provide students with a comfortable testing environment.
- Students are permitted to use mathematics manipulatives during the EMSA if they would ordinarily be permitted to use manipulatives in your classroom.
- No talking or communication between students is permitted during testing.
- Testing administrators should adhere to the guidelines and administration instructions.
- The EMSA, Student Surveys, and reading tests are intended to be read aloud to students by the testing administrator.
- The EMSA administration script indicates that teachers should read the question two times. However, it is permissible for teachers to read the problem more than two times if needed. To ensure uniformity of testing, if the problem is read more than two times, the teacher should read the entire problem to the students and not parts of the problem.

Administering the EMSA, Student Surveys, and Reading Test

It is assumed that the classroom teacher will administer the EMSA, Student Surveys, and reading test; however, other school personnel (such as a paraprofessional or even a substitute teacher) can administer them, providing they follow the assessment protocol as described below. Each packet can be administered on different days or at different times as long as all assessments and surveys are completed during the testing window. Record the testing administration date as well as the start and ending times for both the EMSA test and the Student Surveys on the front page of this administration guide.

The testing conditions should be consistent with the testing conditions for other student assessments administered in the classroom. For example, students should space out the desks or use student “privacy folders” if that is what they would usually do.

Avoid reading problems or answering student questions in a way that may offer clues to the correct answer. Student responses should reflect their current math knowledge. To ensure that the students’ test responses are valid, it is important that appropriate procedures are followed when administering the EMSA, Student Surveys, and reading test. These procedures include:

- Administration of the appropriate test level (Grade 3 assessment for Grade 3 students, etc.)
- Adherence to the guidelines and administration instructions in order to provide a standardized testing protocol across classrooms
- Maintenance of test security

Accommodations

Students with special academic plans (e.g., IEP, 504, ELL) may receive whatever accommodations are specified in their plans.

Testing Time Allocation

Administration of the EMSA, student surveys, and reading test should take no more than 90 minutes combined. These are not timed assessments, and students should be allowed adequate time to answer the questions.

Submitting Materials

Upon conclusion of testing, repack all used and unused EMSA packets (blue paper), Student Survey (green paper), and reading assessment packets in their original envelope. Repack parent permission form packets (yellow paper) and student assent forms in the envelope they came in. Also, please be sure to include the REALM Study Administration Guide (this document) in the envelope. A member of the project team will come collect the testing materials from your school on May 6, 2019. Please have all materials in the front office. You will receive an email prior to pick-up to ensure that all materials are ready in the front office.

If you have questions about this process, contact Charity Bauduin (cbauduin2@lsi.fsu.edu).

Please turn to the next page for the EMSA Administration Script. Instructions for administering the Student Surveys are found on page 17 and for the reading test on page 27.

**Spring 2019 EMSA Administration Instructions
(Blue Booklets)**

This page is intentionally left blank.

EMSA Spring Administration Instructions – Grade 3

[The boxes contain the script that you will read to students.]

You are about to take a math assessment for a research project at Florida State University. You will need a pencil.

Verify that all students have a pencil. Pass out the student assent forms to those students who have forms labeled for them. There are additional blank forms for any new students in your class.

Before we start the math assessment, I am going to read you a letter from scientists at Florida State University. The letter will explain their research project and will ask if you want to be part of it. Some of you have already signed this letter in the fall and will not need to sign it again.

Your teacher is participating in a research study about how teachers teach math, how students feel about math, and how students learn in math. The people doing the research would like to have copies of your answers on some tests that you will take and some questions your teacher will ask you. For some students, we will also observe your classroom on a few days.

If you agree to be part of this research, you will answer questions about how you feel about math and other subjects and also do some math problems and other tests. This will take a total of about 60-90 minutes at three different times. You might think some of the questions are too hard, but that is ok, we just want you to try your best on them. This study may help teachers learn better ways to help students with their math skills. We will also ask your school for your scores on some tests you normally take each year.

As a thank you for helping us out, you will get a small gift.

Your parents have given their permission for you to do this study. But even if your parents gave permission, you can still decide to not do the study and that is ok. You do not have to sign this paper if you do not want the researchers to have this information. Your choice to participate in the research study will not affect your grades in school.

You can ask any questions that you have about this study. If you have questions later, you can contact Dr. Colleen Ganley at cganley@fsu.edu with your questions.

Signing your name at the bottom means that you agree to be in this study and to do the tasks listed above.

If you want to be a part of this research project, write your name on the line and then write today's date which is _____

If you have a question, you can raise your hand.

Address any questions.

Verify that all students have written their first and last name on the assent form. You may help students by writing their name if they give you verbal assent. Occasionally a student does not assent, if this happens, you can let them know that the researchers will not look at their answers. Please collect all assent forms.

I will now pass out the math assessments. The assessments are labeled with your names. When you receive the assessment, keep it face up, and do not turn any pages; we will all begin at the same time after I go over the instructions. Some of these questions may be hard, but don't worry and just try your best.

Ensure that students and pre-labeled test booklets are properly paired (i.e., each student receives the test booklet that has his or her name written on it). If your students would ordinarily be permitted to use manipulatives and/or scratch paper in this type of situation, ensure that they are available at this time or remind them of your policies for how to access them.

Before we start, please write today's date, which is _____.

Check to make sure students are writing the correct date.

Turn the page. You should see Page 1 in the corner. Look at the first example. It says: What grade are you in? The correct answer choice is 3, for third grade. Notice how the bubble below the 3 has been filled in for you. You are going to mark your answer choices the same way, by filling in the bubble below the answer choice you think is correct.

For some problems, you will draw a circle around the answer you think is correct.

Look at the second example. It says, 'Draw a circle around the picture of the car.' The correct answer is circled for you. For some problems, you are going to draw a circle around your answer the same way, by drawing a circle around the answer you think is correct.

For some problems, you will write your answer in a box.

Look at the third example. It says, "Write the number four in the box." The correct answer is written for you in the box. For some problems, you are going to write your answer in the same way, by writing your answer in a box.

For some questions you will be asked to write your answer on a line.

For each problem, I would like you to try hard to figure out which answer is correct. If you are not sure, mark the answer you think is best.

I will read all of the problems to you. Please do not say any answers out loud. You will answer all of the questions by writing on your paper.

You may underline words in the problems if you find that helpful. Also, you can use the blank space on the paper to work out your answers.

Are there any questions?

Address any questions.

I will read each problem two times. If there are no more questions, turn to page 2.

Pause; check to ensure all students are on the correct page.

The first problem is,

Write your answer in the box.

I am going to read the problem one more time:

Write your answer in the box.

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

Turn to page 3.

Pause; check to ensure all students are on the correct page.

Write your answer in the box.

I am going to read the problem one more time:

? Write your answer in the box.

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

Turn to page 4.

Pause; check to ensure all students are on the correct page.

? Write your answer in the box.

I am going to read the problem one more time:

? Write your answer in the box.

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

Turn to page 5.

Pause; check to ensure all students are on the correct page.

Write your answer in the box.

I am going to read the problem one more time:

Write your answer in the box.

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

Turn to page 6.

Pause; check to ensure all students are on the correct page.

Write your answer in the box.

I am going to read the problem one more time:

? Write your answer in the box.

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

Turn to page 7.

Pause; check to ensure all students are on the correct page.

For the next five problems I am going to read an equation aloud to you. If the equation is correct, circle the word yes. If the equation is not correct, circle the word no.

The first equation is . Again, . If the equation is correct, circle the word yes. If the equation is not correct, circle the word no.

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

The last equation on this page is . Again, . If the equation is correct, circle the word yes. If the equation is not correct, circle the word no.

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

Turn to page 8.

Pause; check to ensure all students are on the correct page.

The first equation on this page is [REDACTED]. Again, [REDACTED]. If the equation is correct, circle the word yes. If the equation is not correct, circle the word no.

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

The last equation on this page is [REDACTED]. Again, [REDACTED]. If the equation is correct, circle the word yes. If the equation is not correct, circle the word no.

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

Turn to page 9.

Pause; check to ensure all students are on the correct page.

The equation on this page is [REDACTED]. Again, [REDACTED]. If the equation is correct, circle the word yes. If the equation is not correct, circle the word no.

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

Turn to page 10.

[REDACTED]

I am going to read the problem one more time. [REDACTED]

[REDACTED]

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

Turn to page 11.

Pause; check to ensure all students are on the correct page.

[REDACTED]

Fill in the bubble under the answer you think is correct.

[REDACTED]

Fill in the bubble under the answer you think is correct.

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

Turn to page 12.

Pause; check to ensure all students are on the correct page.

[REDACTED]

I am going to read the problem one more time. [REDACTED]

[REDACTED]

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

Turn to page 13.

Pause; check to ensure all students are on the correct page.

Fill in the bubble under the answer you think is correct.

I am going to read the problem one more time.

Fill in the bubble under the answer you think is correct.

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

Turn to page 14.

Pause; check to ensure all students are on the correct page.

There are two problems on this page. The first problem is,

Write your answer on the line.

I am going to read the problem one more time.

? Write your answer on the line.

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

The last problem on this page is .

Fill in the bubble below the number you think should go in the box.

I am going to read the problem one more time. . Fill in the bubble below the number you think should go in the box.

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

Turn to page 15.

Pause; check to ensure all students are on the correct page.

There are two problems on this page. The first problem is ? Write your answer in the box.

I am going to read the problem one more time. ? Write your answer in the box.

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

The last problem on this page is . Write your answer in the box.

I am going to read the problem one more time. ? Write your answer in the box.

When you are finished, put your pencil down.

Pause and wait for all students to complete the item.

Now you are going to work on some problems on your own. The next four pages have some addition, subtraction, and multiplication problems that you will solve at your own pace. Your job is to find out what number goes in the box to make the equation correct. Write your answer for each problem in the box.

Remember to look closely at the symbol to decide if it is an addition, subtraction, or multiplication problem. When I say “begin” you can start answering the questions. When you get to the end of the first page, continue on to the next few pages until you reach the stop sign at the bottom of the last page. Are there any questions?

Address any questions.

Turn to page 16. You may begin.

Circulate as students work on the problems. Provide students with ample time to complete the problems. Once you see that all students have completed the problems, please end the assessment.

We are all finished with this assessment. I am going to collect all the booklets.

Collect all testing materials.

Appendix I. Scoring Criteria

Tables I.1, I.2, I.3, and I.4 provide information on how items were scored and the data entry procedure for each item. The rubrics used to score the remaining items in the Problem Solving and Modeling section of the test follow the table.

Table I.1. Grade K Scoring Key

Bank Item Number	Item description	Response format	Data entry	Correct response
CT1.4.S1		Selected	Enter the number that corresponds to the position of the student's selected response. Enter DNS for did not solve, UI for unclear intent, CR for constructed response, or MR for multiple responses.	
WP22.S1		Selected	Enter the number that corresponds to the position of the student's selected response. Enter DNS for did not solve, UI for unclear intent, CR for constructed response, or MR for multiple responses.	
WP34.S2		Selected	Enter the number that corresponds to the position of the student's selected response. Enter DNS for did not solve, UI for unclear intent, CR for constructed response, or MR for multiple responses.	
WP24.C		Constructed	Enter exactly as written. Enter UI for unclear intent or NR for no response.	
ES5.S1		Selected	Enter "yes" if the student circled yes. Enter "no" if the student circled no. Enter DNS for did not solve, UI for unclear intent, or MR for multiple responses.	
ES8.S1		Selected	Enter "yes" if the student circled yes. Enter "no" if the student circled no. Enter DNS for did not solve, UI for unclear intent, or MR for multiple responses.	
ES15.S1		Selected	Enter "yes" if the student circled yes. Enter "no" if the student circled no. Enter DNS for did not solve, UI for unclear intent, or MR for multiple responses.	
ES1.3.S1		Selected	Enter "yes" if the student circled yes. Enter "no" if the student circled no. Enter DNS for did not solve, UI for unclear intent, or MR for multiple responses.	
FPW2.2.S1		Selected	Enter "1" for all shapes that were circled, underlined, or otherwise indicated. Enter "0" for all shapes that were not circled or indicated. Enter DNS for did not solve.	
FPW3.3.S1		Selected	Enter "1" for all shapes that were circled, underlined, or otherwise indicated. Enter "0" for all shapes that were not circled or indicated. Enter DNS for did not solve.	
CBF15.S2		Selected	Enter the number that corresponds to the position of the student's selected response. Enter DNS for did not solve, UI for unclear	

			intent, CR for constructed response, or MR for multiple responses.	
CBF21.S2		Selected	Enter the number that corresponds to the position of the student's selected response. Enter DNS for did not solve, UI for unclear intent, CR for constructed response, or MR for multiple responses.	
CBF3.S2		Selected	Enter the number that corresponds to the position of the student's selected response. Enter DNS for did not solve, UI for unclear intent, CR for constructed response, or MR for multiple responses.	
CWN1.1.S2		Selected	Enter the number that corresponds to the position of the student's selected response. Enter DNS for did not solve, UI for unclear intent, CR for constructed response, or MR for multiple responses.	
CWN2.1.S2		Selected	Enter the number that corresponds to the position of the student's selected response. Enter DNS for did not solve, UI for unclear intent, CR for constructed response, or MR for multiple responses.	
CBF35.C		Constructed	Enter exactly as written. Enter UI for unclear intent or NR for no response.	
CWN25.C		Constructed	Enter exactly as written. Enter UI for unclear intent or NR for no response.	

Table I.2. Grade 1 Scoring Key

Bank Item Number	Item description	Response format	Data entry	Correct response
WP34.S2		Selected	Enter the number that corresponds to the position of the student's selected response. Enter DNS for did not solve, UI for unclear intent, CR for constructed response, or MR for multiple responses.	
WP24.C		Constructed	Enter exactly as written. Enter UI for unclear intent or NR for no response.	
WP4.2.C		Constructed	Enter exactly as written. Enter UI for unclear intent or NR for no response.	
WP43.C		Constructed	Enter exactly as written. Enter UI for unclear intent or NR for no response.	
ES5.S1		Selected	Enter "yes" if the student circled yes. Enter "no" if the student circled no. Enter DNS for did not solve, UI for unclear intent, or MR for multiple responses.	
ES8.S1		Selected	Enter "yes" if the student circled yes. Enter "no" if the student circled no. Enter DNS for did not solve, UI for unclear intent, or MR for multiple responses.	
ES15.S1		Selected	Enter "yes" if the student circled yes. Enter "no" if the student circled no. Enter DNS for did not solve, UI for unclear intent, or MR for multiple responses.	
ES1.3.S1		Selected	Enter "yes" if the student circled yes. Enter "no" if the student circled no. Enter DNS for did not solve, UI for unclear intent, or MR for multiple responses.	
ES18.S1		Selected	Enter "yes" if the student circled yes. Enter "no" if the student circled no. Enter DNS for did not solve, UI for unclear intent, or MR for multiple responses.	
FPW2.2.S1		Selected	Enter "1" for all shapes that were circled, underlined, or otherwise indicated. Enter "0" for all shapes that were not circled or indicated. Enter DNS for did not solve.	
FPW3.3.S1		Selected	Enter "1" for all shapes that were circled, underlined, or otherwise indicated. Enter "0" for all shapes that were not circled or indicated. Enter DNS for did not solve.	
FPW3.5.S1		Selected	Enter "1" for all shapes that were circled, underlined, or otherwise indicated. Enter "0" for all shapes that were not circled or indicated. Enter DNS for did not solve.	
FPW14.S1		Selected	Enter the number that corresponds to the position of the student's selected response. Enter DNS for did not solve, UI for unclear	

			intent, CR for constructed response, or MR for multiple responses.	
CBF3.S2		Selected	Enter the number that corresponds to the position of the student's selected response. Enter DNS for did not solve, UI for unclear intent, CR for constructed response, or MR for multiple responses.	
CBF10.S2		Selected	Enter the number that corresponds to the position of the student's selected response. Enter DNS for did not solve, UI for unclear intent, CR for constructed response, or MR for multiple responses.	
CWN2.1.S2		Selected	Enter the number that corresponds to the position of the student's selected response. Enter DNS for did not solve, UI for unclear intent, CR for constructed response, or MR for multiple responses.	
CWN4.1.C		Constructed	Enter exactly as written. Enter UI for unclear intent or NR for no response.	
CBF35.C		Constructed	Enter exactly as written. Enter UI for unclear intent or NR for no response.	
CBF5.C		Constructed	Enter exactly as written. Enter UI for unclear intent or NR for no response.	
ES3.3.C		Constructed	Enter exactly as written. Enter UI for unclear intent or NR for no response.	
CWN25.C		Constructed	Enter exactly as written. Enter UI for unclear intent or NR for no response.	
CTWN6.1.C		Constructed	Enter exactly as written. Enter UI for unclear intent or NR for no response.	

Table I.3. Grade 2 Scoring Key

Bank Item Number	Item description	Response format	Data entry	Correct response
WP4.2.C		Constructed	Enter exactly as written. Enter UI for unclear intent or NR for no response.	
WP26.C		Constructed	Enter exactly as written. Enter UI for unclear intent or NR for no response.	
WP16.C		Constructed	Enter exactly as written. Enter UI for unclear intent or NR for no response.	
WP43.C		Constructed	Enter exactly as written. Enter UI for unclear intent or NR for no response.	
WP55.C		Constructed	Enter exactly as written. Enter UI for unclear intent or NR for no response.	
ES5.S1		Selected	Enter “yes” if the student circled yes. Enter “no” if the student circled no. Enter DNS for did not solve, UI for unclear intent, or MR for multiple responses.	
ES8.S1		Selected	Enter “yes” if the student circled yes. Enter “no” if the student circled no. Enter DNS for did not solve, UI for unclear intent, or MR for multiple responses.	
ES15.S1		Selected	Enter “yes” if the student circled yes. Enter “no” if the student circled no. Enter DNS for did not solve, UI for unclear intent, or MR for multiple responses.	
ES1.3.S1		Selected	Enter “yes” if the student circled yes. Enter “no” if the student circled no. Enter DNS for did not solve, UI for unclear intent, or MR for multiple responses.	
ES18.S1		Selected	Enter “yes” if the student circled yes. Enter “no” if the student circled no. Enter DNS for did not solve, UI for unclear intent, or MR for multiple responses.	
FPW2.2.S1		Selected	Enter “1” for all shapes that were circled, underlined, or otherwise indicated. Enter “0” for all shapes that were not circled or indicated. Enter DNS for did not solve.	

FPW3.3.S1		Selected	Enter “1” for all shapes that were circled, underlined, or otherwise indicated. Enter “0” for all shapes that were not circled or indicated. Enter DNS for did not solve.	
FPW3.5.S1		Selected	Enter “1” for all shapes that were circled, underlined, or otherwise indicated. Enter “0” for all shapes that were not circled or indicated. Enter DNS for did not solve.	
FPW14.S1		Selected	Enter the number that corresponds to the position of the student’s selected response. Enter DNS for did not solve, UI for unclear intent, CR for constructed response, or MR for multiple responses.	
CBF3.S2		Selected	Enter the number that corresponds to the position of the student’s selected response. Enter DNS for did not solve, UI for unclear intent, CR for constructed response, or MR for multiple responses.	
CBF10.S2		Selected	Enter the number that corresponds to the position of the student’s selected response. Enter DNS for did not solve, UI for unclear intent, CR for constructed response, or MR for multiple responses.	
CWN4.1.C		Constructed	Enter exactly as written. Enter UI for unclear intent or NR for no response.	
CWN3.4.C		Constructed	Enter exactly as written. Enter UI for unclear intent or NR for no response.	
CBF35.C		Constructed	Enter exactly as written. Enter UI for unclear intent or NR for no response.	
CBF5.C		Constructed	Enter exactly as written. Enter UI for unclear intent or NR for no response.	
ES3.3.C		Constructed	Enter exactly as written. Enter UI for unclear intent or NR for no response.	
CTWN6.1.C		Constructed	Enter exactly as written. Enter UI for unclear intent or NR for no response.	
CWN42.C		Constructed	Enter exactly as written. Enter UI for unclear intent or NR for no response.	
ES21.C		Constructed	Enter exactly as written. Enter UI for unclear intent or NR for no response.	

CWN33.2.C		Constructed	Enter exactly as written. Enter UI for unclear intent or NR for no response.	

Table I.4. Grade 3 Scoring Key

Bank Item Number	Item description	Response format	Data entry	Correct response
WP26.C		Constructed	Enter exactly as written. Enter UI for unclear intent or NR for no response.	
WP16.C		Constructed	Enter exactly as written. Enter UI for unclear intent or NR for no response.	
WP35.C		Constructed	Enter exactly as written. Enter UI for unclear intent or NR for no response.	
WP43.C		Constructed	Enter exactly as written. Enter UI for unclear intent or NR for no response.	
WP55.C		Constructed	Enter exactly as written. Enter UI for unclear intent or NR for no response.	
ES5.S1		Selected	Enter “yes” if the student circled yes. Enter “no” if the student circled no. Enter DNS for did not solve, UI for unclear intent, or MR for multiple responses.	
ES8.S1		Selected	Enter “yes” if the student circled yes. Enter “no” if the student circled no. Enter DNS for did not solve, UI for unclear intent, or MR for multiple responses.	
ES15.S1		Selected	Enter “yes” if the student circled yes. Enter “no” if the student circled no. Enter DNS for did not solve, UI for unclear intent, or MR for multiple responses.	
ES1.3.S1		Selected	Enter “yes” if the student circled yes. Enter “no” if the student circled no. Enter DNS for did not solve, UI for unclear intent, or MR for multiple responses.	
ES18.S1		Selected	Enter “yes” if the student circled yes. Enter “no” if the student circled no. Enter DNS for did not solve, UI for unclear intent, or MR for multiple responses.	
FPW3.5.S1		Selected	Enter “1” for all shapes that were circled, underlined, or otherwise indicated. Enter “0” for all shapes that were not circled or indicated. Enter DNS for did not solve.	

FPW14.S1		Selected	Enter the number that corresponds to the position of the student's selected response. Enter DNS for did not solve, UI for unclear intent, CR for constructed response, or MR for multiple responses.	
FNL2.1.S1		Selected	Enter the number that corresponds to the position of the student's selected response. Enter DNS for did not solve, UI for unclear intent, CR for constructed response, or MR for multiple responses.	
FWP7.S1		Selected	Enter the number that corresponds to the position of the student's selected response. Enter DNS for did not solve, UI for unclear intent, CR for constructed response, or MR for multiple responses.	
FPW7.C		Constructed	Enter exactly as written. Enter UI for unclear intent or NR for no response.	
CBF10.S2		Selected	Enter the number that corresponds to the position of the student's selected response. Enter DNS for did not solve, UI for unclear intent, CR for constructed response, or MR for multiple responses.	
CWN4.1.C		Constructed	Enter exactly as written. Enter UI for unclear intent or NR for no response.	
CWN3.4.C		Constructed	Enter exactly as written. Enter UI for unclear intent or NR for no response.	
CBF35.C		Constructed	Enter exactly as written. Enter UI for unclear intent or NR for no response.	
ES3.3.C		Constructed	Enter exactly as written. Enter UI for unclear intent or NR for no response.	
CTWN6.1.C		Constructed	Enter exactly as written. Enter UI for unclear intent or NR for no response.	
CWN49.C		Constructed	Enter exactly as written. Enter UI for unclear intent or NR for no response.	
ES.27.C		Constructed	Enter exactly as written. Enter UI for unclear intent or NR for no response.	

CF7.C		Constructed	Enter exactly as written. Enter UI for unclear intent or NR for no response.	
ES21.C		Constructed	Enter exactly as written. Enter UI for unclear intent or NR for no response.	
CWN33.2.C		Constructed	Enter exactly as written. Enter UI for unclear intent or NR for no response.	
CWN50.C		Constructed	Enter exactly as written. Enter UI for unclear intent or NR for no response.	

Appendix J. Within-grade IRT Estimates

J.1. Within-grade Item Parameter Estimates

Table J.1. Descriptive Statistics of Discrimination Estimates and Difficulty Estimates of Each Item in Each Grade-Level Test

	Kindergarten		Grade 1		Grade 2		Grade 3	
	<i>a</i>	<i>b</i>	<i>a</i>	<i>b</i>	<i>a</i>	<i>b</i>	<i>a</i>	<i>b</i>
Mean	1.36	−0.76	1.28	−0.59	1.53	−0.55	1.51	−0.81
St. dev.	0.50	1.15	0.46	1.23	1.11	1.18	0.64	0.88
Minimum	0.42	−3.49	0.49	−2.57	0.42	−2.71	0.73	−2.86
Maximum	2.35	1.22	2.27	1.47	5.52	1.91	3.02	1.22
Skewness	0.08	−0.80	0.39	0.22	2.73	0.20	1.45	−0.17
Kurtosis	0.00	1.22	−0.39	−0.95	8.82	−0.23	1.55	0.70

Note. *a* = item discrimination index; *b* = item difficulty index.

Table J.2. Parameter Estimates and Standard Errors for Final-Scale Items of Kindergarten Test Modeled with 2PL

Bank item number	Final-scale item #	<i>a</i>	<i>s.e.</i>	<i>b</i>	<i>s.e.</i>
CT1.4.S1	1	0.96	0.26	−3.49	0.78
WP22.S1	2	1.42	0.24	−1.99	0.24
WP34.S2	3	0.85	0.13	0.22	0.12
WP24.C	4	1.70	0.20	0.05	0.08
ES15.S1	5	1.05	0.15	−0.48	0.11
ES1.3.S1	6	0.42	0.11	−0.53	0.25
FPW2.2.S1	7	1.09	0.14	0.21	0.10
FPW3.3.S1	8	0.81	0.13	1.22	0.21
CBF15.S2	9	1.65	0.27	−1.91	0.21
CBF21.S2	10	2.35	0.28	−0.70	0.08
CBF3.S2	11	2.02	0.24	−0.71	0.09
CWN1.1.S2	12	1.49	0.24	−1.78	0.21
CWN2.1.S2	13	1.51	0.20	−0.92	0.11
CBF35.C.	14	1.42	0.17	−0.35	0.09
CWN25.C	15	1.65	0.19	−0.22	0.08

Note. Final-scale item # = the newly generated item number after two items (originally numbered 5 and 6) were excluded. *a* = item discrimination index; *b* = item difficulty index.

Table J.3. Parameter Estimates and Standard Errors for Final-Scale Items of Grade 1 Test Modeled with 2PL

Bank item number	Final-scale item #	<i>a</i>	<i>s.e.</i>	<i>b</i>	<i>s.e.</i>
WP34.S2	1	1.74	0.20	–0.39	0.08
WP24.C	2	2.27	0.27	–0.85	0.08
WP4.2.C	3	1.57	0.19	–1.01	0.11
WP43.C	4	1.83	0.23	0.70	0.08
ES15.S1	5	1.07	0.18	–1.79	0.24
ES1.3.S1	6	1.03	0.15	–1.10	0.15
ES18.S1	7	0.92	0.13	–0.21	0.11
FPW2.2.S1	8	1.78	0.21	–1.15	0.11
FPW3.3.S1	9	0.85	0.12	0.45	0.12
FPW3.5.S1	10	0.97	0.15	1.27	0.18
FPW14.S1	11	0.49	0.11	1.47	0.34
CBF3.S2	12	1.08	0.20	–2.57	0.38
CWN2.1.S2	13	1.42	0.22	–1.75	0.19
CWN4.1.C	14	1.73	0.20	–0.59	0.08
CBF35.C	15	0.96	0.17	–2.23	0.32
ES3.3.C	16	1.32	0.16	1.17	0.13
CWN25.C	17	1.30	0.21	–1.86	0.22
CTWN6.1.C	18	0.73	0.12	–0.17	0.13

Note. Final-scale item # = the newly generated item number after three items (originally numbered 5, 6, and 19) were excluded. *a* = item discrimination index; *b* = item difficulty index.

Table J.4. Parameter Estimates and Standard Errors for Final-Scale Items of Grade-2 Test Modeled with 2PL

Bank item number	Final-scale item #	<i>a</i>	<i>s.e.</i>	<i>b</i>	<i>s.e.</i>
WP4.2.C	1	1.47	0.20	-1.47	0.15
WP26.C	2	1.16	0.14	-0.56	0.10
WP16.C	3	1.43	0.16	-0.44	0.09
WP43.C	4	1.31	0.15	-0.13	0.08
WP55.C	5	1.17	0.15	1.33	0.15
ES15.S1	6	1.33	0.23	-2.13	0.26
ES1.3.S1	7	1.27	0.16	-1.59	0.17
ES18.S1	8	2.12	0.23	-0.32	0.07
FPW2.2.S1	9	1.18	0.17	-1.97	0.22
FPW3.3.S1	10	0.69	0.11	-0.38	0.14
FPW3.5.S1	11	0.62	0.12	1.91	0.36
FPW14.S1	12	0.42	0.10	0.71	0.26
CBF3.S2	13	1.43	0.26	-2.71	0.35
CWN4.1.C	14	1.97	0.24	-1.56	0.14
CWN3.4.C	15	1.35	0.18	-1.35	0.16
ES3.3.C	16	3.19	0.36	0.38	0.06
CTWN6.1.C	17	0.91	0.13	-0.75	0.14
CWN42.C	18	1.36	0.16	-0.25	0.08
ES21.C	19	5.52	0.93	0.62	0.06
CWN33.2.C	20	0.76	0.11	-0.38	0.13

Note. Final-scale item # = the newly generated item number after five items (originally numbered 6, 7, 16, 19, and 20) were excluded. *a* = item discrimination index; *b* = item difficulty index.

Table J.5. Parameter Estimates and Standard Errors for Final-Scale Items of Grade-3 Test Modeled with 2PL

Bank item number	Final-scale item #	<i>a</i>	<i>s.e.</i>	<i>b</i>	<i>s.e.</i>
WP26.C	1	1.52	0.22	-2.04	0.21
WP16.C	2	1.82	0.19	-1.05	0.10
WP35.C	3	1.33	0.15	-0.30	0.08
WP43.C	4	1.53	0.17	-0.42	0.08
WP55.C	5	1.06	0.14	1.22	0.15
ES15.S1	6	1.41	0.28	-2.86	0.40
ES1.3.S1	7	1.24	0.20	-2.07	0.24
ES18.S1	8	1.81	0.19	-0.79	0.08
FPW3.5.S1	9	1.05	0.12	-0.29	0.10
FPW14.S1	10	0.73	0.11	-0.40	0.13
FNL2.1.S1	11	1.00	0.13	-0.92	0.13
FWP7.S1	12	1.13	0.13	-0.88	0.12
FWP7.C	13	1.22	0.14	-1.15	0.13
CWN4.1.C	14	1.86	0.23	-1.78	0.15
CWN3.4.C	15	1.56	0.18	-1.58	0.15
ES3.3.C	16	3.02	0.30	-0.01	0.06
CTWN6.1.C	17	0.99	0.14	-1.43	0.20
CWN42.C	18	1.52	0.20	-1.10	0.12
CWN49.C	19	1.37	0.15	0.08	0.08
ES.27.C	20	3.02	0.30	-0.44	0.06
CF7.C	21	0.84	0.11	-0.63	0.13
ES21.C	22	2.88	0.28	0.09	0.06
CWN33.2.C	23	0.97	0.12	-0.91	0.14
CWN50.C	24	1.39	0.15	0.10	0.08

Note. Final-scale item # = the newly generated item number after four items (originally numbered 6, 7, 16, and 19) were excluded. *a* = item discrimination index; *b* = item difficulty index.

Figure J.1. Within-grade item discrimination estimates for kindergarten test.

Figure J.2. Within-grade item difficulty estimates for kindergarten test.

Figure J.3. Within-grade item discrimination estimates for grade-1 test.

Figure J.4. Within-grade item difficulty estimates for grade-1 test.

Figure J.5. Within-grade item discrimination estimates for grade-2 test.

Figure J.6. Within-grade item difficulty estimates for grade-2 test.

Figure J.7. Within-grade item discrimination estimates for grade-3 test.

Figure J.8. Within-grade item difficulty estimates for grade-3 test.

J.2. Within-grade Test Information and Estimated Latent Trait

Figure J.9. Within-grade test information curve and conditional standard error of measurement (CSEM) for the final scale format of kindergarten test.

Figure J.10. Within-grade test information curve and CSEM for the final scale format of grade-1 test (base scale).

Figure J.11. Within-grade test information curve and CSEM for the final scale format of grade-2 test.

Figure J.12. Within-grade test information curve and CSEM for the final scale format of grade-3 test.

Figure J.13. Within-grade maximum likelihood estimation (MLE)-based latent estimates for kindergarten students.

Figure J.14. Within-grade expected a posteriori-based latent estimates of kindergarten students.

Figure J.15. Within-grade MLE-based latent estimates of grade-1 students.

Figure J.16. Within-grade EAP-based latent estimates of grade-1 students.

Figure J.17. Within-grade MLE-based latent estimates of grade-2 students.

Figure J.18. Within-grade EAP-based latent estimates of grade-2 students.

Figure J.19. Within-grade MLE-based latent estimates of grade-3 students.

Figure J.20. Within-grade EAP-based latent estimates of grade-3 students.

Appendix K. EMSA Spring 2018 Sample

Originally, in the EMSA 2018 data, 109 kindergarten students participated and the test included 18 items; in grade 1, 55 students and 22 items; in grade 2, 37 students and 26 items; and in grade 3, 46 students and 29 items.

For the EMSA 2018 missing-data analysis, we followed the same rule as for the EMSA 2019. After missing data were excluded, 108 students remained in the kindergarten group, 55 in grade 1, 37 in grade 2, and 46 in grade 3.

For scoring of the EMSA 2018 samples, the items that did not exist in the EMSA 2019 tests were also dropped from the EMSA 2018 tests because of the small sample sizes in each grade level of the latter. In the final scale of the EMSA 2018 tests, 17 items remained in the kindergarten test, 19 in the grade-1 test, 22 in the grade-2 test, and 19 in the grade-3 test.

The observed total test score calculated at each grade level was used to depict the distribution. Figures K.1–K.4 illustrate the minimum and maximum of the total test score for each grade level as well as the percentages of students at the observed test score. The kindergarten observed score distribution in Figure M.1 shows that none of the students got the minimum score of 0, whereas 4 scored perfectly (mean 10.60, standard deviation 3.56). Likewise, on the grade-1 test, none of the students got a score of 0, whereas 2 scored perfectly (mean 13.71, standard deviation 3.51). On the grade-2 test, no student got either the minimum or the maximum score (mean 13.73, standard deviation 4.60). Likewise, on the grade-3 test, no student either the minimum or the maximum score (mean 9.78, standard deviation 4.68).

Figure K.1. The distribution of observed test scores for the kindergarten test.

Figure K.2. Distribution of observed test scores for the grade-1 test.

Figure K.3. Distribution of observed test scores for the grade-2 test.

Figure K.4. Distribution of observed test scores for the grade-3 test.

K.1. Item Response Theory Model

Items in each grade level test were dichotomously scored, either constructed response or selected response/multiple choice items. Because our purpose was to obtain latent trait scores for the EMSA 2018 samples through the fixed item parameter calibration approach, the two-parameter logistic model was chosen as it was for the EMSA 2019 samples. Equation 2 in Section 7 shows the formula of the two-parameter logistic model. Below, we briefly explain how the latent trait scores at each grade level were obtained by means of the fixed item parameter calibration approach.

K.2. Person Scoring and Test Information Function

Given that the EMSA 2018 sample sizes were small but assumed to come from the same population as did EMSA 2019 samples, we used item parameter estimates from the EMSA 2019 tests' results to obtain latent trait scores for the EMSA 2018 samples. The EMSA 2018 and EMSA 2019 tests were similar in having common items across the same grade levels. For example, the EMSA 2018 kindergarten test consisted of 17 items, 15 of which were also included in the EMSA 2019 kindergarten test. As a result, two unique items had to be estimated in the EMSA 2018 kindergarten test, but these two items were in fact administered in the EMSA 2019 grade-1 test. Because we had already put item parameter estimates in the EMSA 2019 tests on a common scale, we used the estimates from the EMSA 2019 grade-1 test for those two items.

Tables K.1–K.4 show the common-item match with the EMSA 2019 tests for each grade level and the corresponding item parameter estimates. Overall, by fixing the item parameter estimates of the EMSA 2018 tests from the EMSA 2019 tests, we scored the students in the EMSA 2018 tests. The scoring was

carried out by flexMIRT 3.5 (Cai, 2017). In addition, the population mean and variance at each grade level were fixed at the corresponding value from the EMSA 2019 tests' results. Figures K.5-K.8 present the relationship between the test information function and the conditional standard error of measurement for each grade level. Figures K.9-K.16 show the distributions of the equated latent trait estimates by both maximum likelihood estimation and expected *a posteriori* method for each grade level.

Table K.1. Common-item match for kindergarten tests

Final-scale item in EMSA 2018 #	Corresponding final-scale item in EMSA 2019	Corresponding <i>a</i> value	Corresponding <i>d</i> value
1	1	0.95	4.34
2	2	1.45	4.38
3	3	1.74	0.67
4	4	2.27	1.94
5	G1.3	1.57	1.58
6	5	1.07	1.92
7	6	1.03	1.14
8	7	1.78	2.05
9	8	0.85	-0.38
10	9	1.69	4.95
11	10	2.07	3.70
12	11	1.08	2.77
13	12	1.59	4.41
14	13	1.42	2.50
15	14	0.96	2.15
16	G1.20	1.32	-1.55
17	15	1.30	2.43

Note. *a* = item discrimination index; *d* = intercept parameter; G1.3 and G1.20 represent 3rd and 12th items in the EMSA Spring 2019 grade-1 test, respectively.

Table K.2. Common-item match for grade-1 tests

Final-scale item in EMSA 2018 #	Corresponding final-scale item in EMSA 2019	Corresponding a value	Corresponding d value
1	1	1.74	0.67
2	2	2.27	1.94
3	3	1.57	1.58
4	4	1.83	-1.28
5	5	1.07	1.92
6	6	1.03	1.14
7	7	0.92	0.20
8	8	1.78	2.05
9	9	0.85	-0.38
10	10	0.97	-1.23
11	11	0.49	-0.72
12	GK.12	2.07	3.70
13	12	1.08	2.77
14	13	1.42	2.50
15	14	1.73	1.02
16	15	0.96	2.15
17	16	1.32	-1.55
18	17	1.30	2.43
19	18	0.73	0.12

Note. a = item discrimination index; d = intercept parameter; GK.12 represents 12th item in the EMSA Spring 2019 kindergarten test.

Table K.3. Common-item match for grade-2 tests

Final-scale item in EMSA 2018 #	Corresponding final-scale item in EMSA 2019	Corresponding a value	Corresponding d value
1	G1.2	2.27	1.94
2	1	1.57	1.58
3	2	1.24	-0.12
4	3	1.50	-0.31
5	G3.3	1.11	-1.09
6	4	1.83	-1.28
7	5	1.20	-2.36
8	6	1.07	1.92
9	7	1.03	1.14
10	8	0.92	0.20
11	9	1.78	2.05
12	10	0.85	-0.38
13	11	0.97	-1.23
14	12	0.49	-0.72
15	13	1.08	2.77
16	14	1.73	1.02
17	15	1.31	0.99
18	16	0.73	0.12
19	17	1.32	-1.55
20	18	1.42	-0.55
21	19	3.09	-4.17
22	20	0.79	-0.21

Note. a = item discrimination index; d = intercept parameter; G1.2 represents 2nd item in the EMSA Spring 2019 grade-1 test; G3.3 represents 3rd in the EMSA Spring 2019 grade-3 test.

Table K.4. Common-item match for grade-3 tests

Final-scale item in EMSA 2018 #	Corresponding final-scale item in EMSA 2019	Corresponding a value	Corresponding d value
1	1	1.29	1.44
2	2	1.48	-0.07
3	3	1.11	-1.09
4	4	1.83	-1.28
5	5	0.86	-2.46
6	6	1.07	1.92
7	7	1.03	1.14
8	8	0.92	0.20
9	9	0.97	-1.23
10	10	0.49	-0.72
11	11	0.81	-0.17
12	12	0.94	-0.26
13	14	1.73	1.02
14	15	1.31	0.75
15	17	0.73	0.12
16	16	1.32	-1.55
17	18	1.27	0.00
18	22	1.97	-2.90
19	23	0.78	-0.18

Note. a = item discrimination index; d = intercept parameter.

Figure K.5. Test information curve and CSEM for the final scale format of kindergarten test.

Figure K.6. Test information curve and CSEM for the final scale format of grade-1 test.

Figure K.7. Test information curve and CSEM for the final scale format of grade-2 test.

Figure K.8. Test information curve and CSEM for the final scale format of grade-3 test.

Figure K.9. Maximum-likelihood estimation (MLE)-based latent estimates for kindergarten students.

Figure K.10. Expected a posteriori (EAP)–based latent estimates for kindergarten students.

Figure K.11. MLE-based latent estimates for grade-1 students.

Figure K.12. EAP-based latent estimates for grade-1 students.

Figure K.13. MLE-based latent estimates for grade-2 students.

Figure K.14. EAP-based latent estimates for grade-2 students.

Figure K.15. MLE-based latent estimates for grade-3 students.

Figure K.16. EAP-based latent estimates for grade-3 students.

Appendix L. Most Common Incorrect Responses for Each Item

Table L.1. Percentage of Grade K Student Responses by Item

Item	Item description	Response format	Correct response	Most frequent incorrect responses			
			Response (%)	Response (%)	Response (%)	Response (%)	Response (%)
CT1.4.S1		SR	(94.6)	(2.8)	(1.4)	(1.0)	(0.2)
WP22.S1		SR	(88.8)	(3.7)	(3.5)	(2.1)	(1.2)
WP34.S2		SR	(45.5)	(20.1)	(18.0)	(14.1)	(1.2)
WP24.C		CR	(47.8)	(11.7)	(11.0)	(4.4)	(3.3)
ES5.S1		SR	(92.5)	(5.2)	(2.3)	–	–
ES8.S1		SR	(92.8)	(4.5)	(2.6)	–	–
ES15.S1		SR	(59.7)	(39.3)	(1.0)	–	–
ES1.3.S1		SR	(55.0)	(42.8)	(1.9)	(0.2)	(0.2)
FPW2.2.S1		SR	(45.0)*	–	–	–	–
FPW3.3.S1		SR	(29.3)*	–	–	–	–
CBF15.S2		SR	(90.1)	(5.9)	(1.7)	(1.2)	(0.5)
CBF21.S2		SR	(70.2)	(17.8)	(7.0)	(3.1)	(1.4)
CBF3.S2		SR	(69.6)	(14.8)	(7.7)	(4.4)	(2.4)
CWN1.1.S2		SR	(87.4)	(6.5)	(3.0)	(1.9)	(1.2)
CWN2.1.S2		SR	(72.3)	(15.7)	(8.6)	(2.3)	(1.0)
CBF35.C		CR	(58.3)	(10.1)	(5.6)	(3.5)	(2.8)
CWN25.C		CR	(55.5)	(4.2)	(3.7)	(3.5)	(3.1)

Note. $n = 573$ valid grade GK tests conducted. Boldface items were removed as a result of initial screening. Items that were not answered were recorded as “NR.” Item responses that were unclear were recorded as “UI.” Items with asterisks were entered as correct or incorrect.

Table L.2. Percentage of Grade 1 Student Responses by Item

Item	Item description	Response format	Correct response	Most frequent incorrect responses			
			Response (%)	Response (%)	Response (%)	Response (%)	Response (%)
WP34.S2		SR	(60.8)	(22.7)	(9.2)	(6.9)	(0.3)
WP24.C		CR	(75.1)	(6.7)	(3.3)	(2.4)	(2.0)
WP4.2.C		CR	(74.9)	(9.4)	(3.1)	(1.9)	(1.4)
WP43.C		CR	(30.4)	(24.1)	(5.5)	(2.8)	(2.5)
ES5.S1		SR	(98.9)	(0.8)	(0.3)	–	–
ES8.S1		SR	(99.1)	(0.8)	(0.2)	–	–
ES15.S1		SR	(82.8)	(16.8)	(0.3)	(0.2)	–
ES1.3.S1		SR	(71.6)	(27.3)	(0.6)	(0.3)	(0.2)
ES18.S1		SR	(53.9)	(45.1)	(0.2)	–	–
FPW2.2.S1		SR	(79.6)*	–	–	–	–
FPW3.3.S1		SR	(41.8)*	–	–	–	–
FPW3.5.S1		SR	(26.5)*	–	–	–	–
FPW14.S1		SR	(33.4)	(39.2)	(27.1)	(0.3)	–
CBF3.S2		SR	(91.2)	(3.6)	(3.1)	(1.6)	(0.3)
CBF10.S2		SR	(83.7)	(7.5)	(5.0)	(2.8)	(0.8)
CWN2.1.S2		SR	(86.8)	(6.7)	(3.4)	(1.9)	(0.8)
CWN4.1.C		CR	(66.3)	(2.0)	(1.9)	(1.9)	(1.7)
CBF35.C		CR	(86.4)	(5.6)	(2.7)	(1.4)	(0.6)
CBF5.C		CR	(83.7)	(3.3)	(3.0)	(2.8)	(1.7)
ES3.3.C		CR	(23.5)	(42.3)	(15.7)	(4.4)	(2.7)
CWN25.C		CR	(86.8)	(2.5)	(1.9)	(1.3)	(1.1)
CTWN6.1.C		CR	(52.7)	(10.0)	(4.7)	(4.5)	(3.4)

Note. $n = 638$ valid grade 1 tests conducted. Boldface items were removed as a result of initial screening. Items that were not answered were recorded as “NR.” Item responses that were unclear were recorded as “UI.” Items with asterisks were entered as correct or incorrect.

Table L.3. Percentage of Grade 2 Student Responses by Item

Item	Item description	Response format	Correct response	Most frequent incorrect responses				
			Response (%)	Response (%)	Response (%)	Response (%)	Response (%)	Response (%)
WP4.2.C		CR	(82.6)	(7.6)	(1.1)	(0.9)	(0.8)	
WP26.C		CR	(62.3)	(12.1)	(4.1)	(3.5)	(2.9)	
WP16.C		CR	(60.9)	(8.2)	(4.7)	(4.1)	(3.2)	
WP43.C		CR	(53.0)	(12.2)	(4.0)	(3.1)	(2.6)	
WP55.C		CR	(22.1)	(4.9)	(4.3)	(3.7)	(2.6)	
ES5.S1		SR	(98.8)	(0.6)	(0.7)	–	–	
ES8.S1		SR	(99.4)	(0.5)	(0.2)	–	–	
ES15.S1		SR	(89.8)	(9.8)	(.5)	–	–	
ES1.3.S1		SR	(82.4)	(16.8)	(0.8)	–	–	
ES18.S1		SR	(59.8)	(39.3)	(0.8)	–	–	
FPW2.2.S1		SR	(86.4)*	–	–	–	–	
FPW3.3.S1		SR	(55.4)*	–	–	–	–	
FPW3.5.S1		SR	(25.0)*	–	–	–	–	
FPW14.S1		SR	(42.7)	(28.2)	(28.1)	(0.6)	(0.2)	
CBF3.S2		SR	(94.8)	(1.7)	(1.5)	(0.9)	(1.1)	
CBF10.S2		SR	(90.8)	(3.8)	(2.7)	(1.2)	(1.2)	
CWN4.1.C		CR	(87.2)	(1.4)	(1.2)	(1.1)	(0.9)	
CWN3.4.C		CR	(79.4)	(4.3)	(2.9)	(1.5)	(1.2)	
CBF35.C		CR	(92.5)	(2.7)	(1.1)	(0.9)	(0.6)	
CBF5.C		CR	(93.1)	(1.2)	(0.9)	(0.8)	(0.6)	
ES3.3.C		CR	(36.3)	(44.1)	(10.4)	(2.7)	(1.1)	
CTWN6.1.C		CR	(63.7)	(8.5)	(6.0)	(4.6)	(2.4)	
CWN42.C		CR	(56.0)	(26.3)	(2.4)	(2.0)	(1.4)	
ES21.C		CR	(26.9)	(34.2)	(4.7)	(3.7)	(3.2)	
CWN33.2.C		CR	(56.0)	(4.6)	(3.5)	(3.2)	(2.6)	

Note. n = 655 valid grade 2 tests conducted. Boldface items were removed as a result of initial screening. Items that were not answered were recorded as “NR.” Item responses that were unclear were recorded as “UI.” Items with asterisks were entered as correct or incorrect.

Table L.4. Percentage of Grade 3 Student Responses by Item

Item	Item description	Response format	Correct response	Most frequent incorrect responses			
			Response (%)	Response (%)	Response (%)	Response (%)	Response (%)
WP26.C		CR	(90.7)	(4.3)	(1.0)	(0.7)	(0.6)
WP16.C		CR	(77.7)	(7.0)	(2.6)	(1.4)	(1.3)
WP35.C		CR	(57.2)	(5.2)	(5.1)	(2.3)	(2.2)
WP43.C		CR	(60.9)	(10.4)	(3.3)	(2.8)	(2.0)
WP55.C		CR	(25.7)	(8.3)	(4.8)	(4.2)	(2.8)
ES5.S1		SR	(99.0)	(0.6)	(0.4)	–	–
ES8.S1		SR	(99.3)	(0.7)	–	–	–
ES15.S1		SR	(95.5)	(3.6)	(0.9)	–	–
ES1.3.S1		SR	(88.1)	(11.1)	(0.7)	–	–
ES18.S1		SR	(71.4)	(28.0)	(0.6)	–	–
FPW3.5.S1		SR	(56.1)*	–	–	–	–
FPW14.S1		SR	(56.1)	(23.6)	(19.1)	(1.0)	(0.1)
FNL2.1.S1		SR	(68.0)	(22.0)	(5.4)	(3.9)	(0.7)
FWP7.S1		SR	(68.4)	(19.7)	(5.8)	(4.5)	(1.6)
FPW7.C		CR	(74.6)	(5.4)	(3.2)	(2.5)	(2.3)
CBF10.S2		SR	(94.3)	(2.0)	(2.0)	(1.3)	(0.1)
CWN4.1.C		CR	(89.6)	(1.4)	(1.2)	(0.9)	(0.9)
CWN3.4.C		CR	(84.9)	(5.2)	(2.5)	(1.0)	(0.7)
CBF35.C		CR	(93.6)	(2.3)	(0.9)	(0.6)	(0.4)
ES3.3.C		CR	(50.3)	(35.2)	(8.7)	(1.0)	(0.7)
CTWN6.1.C		CR	(75.9)	(6.1)	(4.1)	(3.8)	(1.6)
CWN42.C		CR	(76.4)	(15.7)	(1.2)	(0.7)	(0.6)
CWN49.C		CR	(47.7)	(7.1)	(6.5)	(4.2)	(3.5)
ES.27.C		CR	(64.5)	(11.4)	(6.4)	(3.2)	(2.3)
CF7.C		CR	(60.7)	(10.9)	(5.8)	(2.5)	(2.3)
ES21.C		CR	(46.7)	(26.5)	(5.9)	(2.5)	(1.9)
CWN33.2.C		CR	(67.1)	(5.4)	(2.0)	(1.4)	(1.0)
CWN50.C		CR	(47.4)	(5.1)	(3.8)	(2.9)	(2.8)

Note. n = 690 valid grade 3 tests conducted. Boldface items were removed as a result of initial screening. Items that were not answered were recorded as “NR.” Item responses that were unclear were recorded as “UI.” Items with asterisks were entered as correct or incorrect.