Ready for Kindergarten: Kindergarten Readiness Assessment Technical Report Addendum

Fall 2015

Prepared for the Maryland State Department of Education
and the Ohio Department of Education by WestEd

The Ready for Kindergarten: Early Childhood Comprehensive Assessment System is a partnership between the Maryland State Department of Education and the Ohio Department of Education, in collaboration with the Johns Hopkins University Center for Technology in Education and WestEd, that is supported by a Race to the Top–Early Learning Challenge grant from the U.S. Department of Education and the U.S. Department of Health and Human Services (CFDA 84.412A) and by a Race to the Top grant from the U.S. Department of Education (CFDA 84.395).

1 Overview

1.1 Race to the Top—Early Learning Challenge Grants

On December 16, 2011, Maryland and Ohio were each awarded Race to the Top Early Learning Challenge (RTT-ELC) Grants for four years. Along with other projects, these funds supported an innovative partnership between Maryland and Ohio to develop the Ready for Kindergarten Early Childhood Comprehensive Assessment System, which consists of the Kindergarten Readiness Assessment (KRA) and the (formative) Early Learning Assessment. A number of partners have played a vital role in executing Maryland's and Ohio's shared vision for improving kindergarten readiness and early childhood assessments. These partners include the Johns Hopkins University Center for Technology in Education (JHU CTE), WestEd (the Standards, Assessment, and Accountability Services [SAAS] program and the Center for Child & Family Studies [CCFS]), state advisory councils, and a national technical advisory committee (TAC), facilitated by the Council of Chief State School Officers (CCSSO). An updated list of project members from each organization is provided in Appendix A.

1.2 Purpose of the Kindergarten Readiness Assessment (KRA)

The purpose of the KRA is to provide information to stakeholders at the local, regional, and state levels about how well prepared children are for kindergarten. State, district, and school leaders use the KRA to learn about children's levels of preparedness and readiness for kindergarten, which enables programmatic decision-making at the school, district, and state levels. Score information by domain and overall readiness can be summarized by demographic characteristics, in order to pinpoint where there are achievement gaps upon kindergarten entry; how children's prior education and care experiences impact readiness; and where to target resources to better support children identified as at-risk through academic, health, and behavioral supports and interventions. By making aggregated assessment reports available in the Ready for Kindergarten Online system at the individual, classroom, school, and district levels, and by facilitating the integration of the KRA results into longitudinal data systems at the state level, the KRA informs these policy, research, and education decisions. Families and teachers learn about each child's skills, learning, and developmental needs so that the teachers can identify strengths and weaknesses for each child.

1.3 Fall 2015 Administration Window

The KRA was administered statewide in Maryland and Ohio between the first day of school and November 1, 2015.

1.4 Purpose of This Report

The fall 2015 administration of the KRA incorporated enhancements and improvements to the technology and professional development components. It was also the first administration of the KRA in its reduced length.

This report is an addendum to the *KRA Technical Report—Fall 2014* and provides the technical qualities of the KRA based on the administration in fall 2015.

2 KRA Design

2.1 KRA Item Types

A KRA item is one question or observation that aligns to a specific essential skill and knowledge statement from within the Common Language Standards and that results in one recorded score. In some instances, multiple items are clustered around a common stimulus (e.g., a story), but multiple item scores are recorded (one for each item in the cluster). The KRA comprises three item types: selected response, performance task, and observational rubric.

2.2 KRA Blueprint

The KRA Blueprint outlines the distribution of items by type, total items, total raw points, and percentage of total raw points across the domains defined in the Common Language Standards. Table 2.2 provides the distribution of item types across the domains in KRA 1.5.

Table 2.2—KRA 1.5 Blueprint

Domain	Selected Response	Performance Task	Observational Rubric	Total Items	Total Raw Points	Percentage of Total Raw Points
Language and Literacy	6	9	2	17	34	35
Mathematics	3	11	0	14	25	26
Physical Well-Being & Motor Development	0	0	7	7	14	14
Social Foundations	0	0	12	12	24	25
Total	9	20	21	50	97	100

3 KRA Data Analyses, Standard Setting, and Reporting

3.1 Item Scores

The KRA 1.5 included dichotomous and polytomous items. All selected-response items were dichotomous and were scored 0–1. Performance-task items were either dichotomous or polytomous and were scored 0–1, 0–2, or 0–3. Observational-rubric items were polytomous and were scored 0–2.

KRA items were scored by the teacher who administered the assessment, unless a student completed items via the Ready for Kindergarten app. Items administered via the Ready for Kindergarten app were auto-scored.

In some circumstances, a teacher determined that an item could not be administered to a student after following the processes described in the *Guidelines on Allowable Supports for Administration of the KRA* document. In this scenario, the teacher entered “Not Scorable” for that particular item.

3.2 Data Analyses

3.2.1 Classical Item Analysis

All KRA items were evaluated for their mean, standard deviation, difficulty (*p*-value), and discrimination (item-rest correlation). These statistics for the fall 2015 administration are provided in Appendix B.

Table 3.2.1 provides a summary of the *p*-values, item-rest correlations, ranges, and Cronbach's alphas from the fall 2015 administration.

Table 3.2.1—Summary of *p*-Values, Item-Rest Correlations, Ranges, and Cronbach's Alpha for the Fall 2015 Administration

	Number of Items	*p*-Value			Item-Rest Correlation			Cronbach's Alpha
		Mean	SD	Range	Mean	SD	Range	
Overall	50	0.75	0.12	0.44–0.93	0.46	0.13	0.23–0.68	0.93
Language and Literacy	17	0.75	0.13	0.57–0.93	0.42	0.11	0.25–0.54	0.81
Mathematics	14	0.69	0.15	0.44–0.92	0.40	0.08	0.26–0.53	0.77
Physical Well-Being and Motor Development	7	0.84	0.04	0.79–0.88	0.56	0.03	0.52–0.60	0.81
Social Foundations	12	0.77	0.07	0.65–0.86	0.66	0.06	0.54–0.73	0.91

3.2.2 IRT Rasch Scaling

After the fall 2014 administration, a one-parameter (i.e., Rasch) item response theory (IRT) model was used to define the relationship between the assumed latent trait (readiness for kindergarten) and the probability of a student correctly answering a given KRA item. WINSTEPS software was used to produce the overall scale and item parameters. The IRT parameters from the fall 2014 administration were retained for the fall 2015 administration.

3.2.3 Ready for Kindergarten App

For the fall 2015 administration, there were 17 items that could be administered via the Ready for Kindergarten app. Administration of items via the app was optional. Table 3.2.3.A shows the number of students who completed items via the app, total number of students, and percentage of the total of students from each state who completed items via the app.

Table 3.2.3.A—Students Administered Items via Ready for Kindergarten App

	Students*	Total	Percentage of Total
Maryland	47,558	65,088	73.1%
Ohio	39,822	123,030	32.4%
Total	87,380	188,118	46.5%

**Includes any student who completed at least one item via the app.*

The fall 2015 administration of the KRA resulted in an 84% increase (64% increase in Maryland and 108% increase in Ohio) in the use of the Ready for Kindergarten App when compared to the fall 2014 administration.

Given the greater number of students who completed items via the Ready for Kindergarten App during the fall 2015 administration, a comparison of the data by mode of administration was completed to determine if the pattern of student responses and scaled scores differed. Caution must be taken when comparing the results of this analysis because the groups of students being compared (i.e., students who were not administered items via the app vs. students who were administered items via the app) are composed of different students. Table 3.2.3.B summarizes the student demographics by the number of items administered via the app.

Table 3.2.3.B—Student Demographics by the Number of Items Administered via the App

		No Items N = 100738		Some Items (1–9 Items) N = 6471		Some Items (10–16 Items) N = 14685		All Items (17 Items) N = 66224	
		n	%	n	%	n	%	n	%
Gender	Female	47,136	47.94	3,065	48.24	6,675	48.75	30,216	48.90
	Male	51,195	52.06	3,289	51.76	7,017	51.25	31,572	51.10
English learner	No	96,352	95.65	5,938	91.76	13,158	89.60	60,770	91.76
	Yes	4,386	4.35	533	8.24	1,527	10.40	5,454	8.24
Special Education	No	92,850	92.17	5,997	92.68	13,666	93.06	61,446	92.79
	Yes	7,888	7.83	474	7.32	1,019	6.94	4,778	7.21
Low SES	No	88,993	88.34	5,423	83.80	10,756	73.24	50,096	75.65
	Yes	11,745	11.66	1,048	16.20	3,929	26.76	16,128	24.35
		Mean	SD	Mean	SD	Mean	SD	Mean	SD
Age		5.74	0.49	5.71	0.40	5.70	0.39	5.71	0.41

In general, students who were administered items via the app performed similarly to the students who were not administered items via the app. Again, caution must be taken when comparing results since the groups of students being compared are composed of different students. Table 3.2.3.C summarizes student results by performance level and number of items administered via the app.

Table 3.2.3.C—Student Results by Performance Level and Number of Items Administered via the App

	No Items		Some Items (1-9 Items)		Some Items (10-16 Items)		All Items (17 Items)	
	n	%	n	%	n	%	n	%
Demonstrating	38,110	41.24	2,559	40.61	6,074	42.91	27,602	42.56
Approaching	32,997	35.71	2,378	37.73	5,349	37.79	25,252	38.94
Emerging	21,301	23.05	1,365	21.66	2,733	19.31	12,000	18.50

In general, the students who were administered app items scored slightly higher, on average, than students who were not administered app items. The mathematics domain was the only exception to this pattern. Table 3.2.3.D provides a summary of student scores (overall and by domain) and number of items administered via the app.

Table 3.2.3.D—Student Scores and Number of Items Administered via the App

	No Items		Some Items (1-9 Items)		Some Items (10-16 Items)		All Items (17 Items)	
	Mean	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
Overall	266.89	14.24	266.81	12.54	267.64	12.09	267.62	11.62
Social Foundations	271.03	20.76	271.71	19.86	273.85	19.63	273.66	19.26
Language and Literacy	265.90	14.75	266.68	13.71	267.87	13.26	267.83	12.42
Mathematics	266.64	15.85	265.01	13.08	264.48	11.80	264.86	11.62
Physical Well-Being and Motor Development	269.82	18.33	270.71	17.40	272.19	16.85	271.90	16.75

A comparison of item-level classical statistics resulted in the identification of four items in which the mode of administration may have an effect on student performance. The four items are in the mathematics domain: MA.1.1.G_A117, MA.3.1.B_A123, MA.3.2.B_A174, MA.4.1.A_A177. One item (MA.1.1.G_A117) exhibits statistics that indicate students who are administered the item via the app perform better. The other three items exhibit statistics that indicate students who are not administered the items via the app perform better. These results may describe the lower mean score in the mathematics domain for students who were administered items via the app, as seen in Table 3.2.3.D above. The comparison of item-level classical statistics by KRA App use is provided in Appendix C.

3.3 Scaled Scores and Reporting

The overall scaled score determines each student's performance level. Domain scaled scores are determined using the same parameters as established for the items when evaluated as an overall test in order to show relative strengths in each student's performance.

Table 3.4.1 summarizes the aggregate distribution of student results by performance level based on the data from both states for the fall 2015 administration. These values are based on only students with complete data. Appendix D provides a more detailed display of the distributions of scaled scores.

Table 3.4—Aggregate Distribution by Performance Level for the Fall 2015 Administration

Reporting Category	Overall Scaled Score Range	Students in Reporting Category (n = 177,720)
Demonstrating	270–298	41.8%
Approaching	258–269	37.1%
Emerging	202–257	21.0%

3.4 Standard Setting Validation

3.4.1 Overview

The purpose of the KRA is to measure students' readiness to engage with kindergarten instruction at the start of school. Therefore, the focus of the performance level descriptors (PLDs) was placed on whether students demonstrate the skills and behaviors that reflect their readiness to engage in instruction based on kindergarten content standards.

- **Demonstrating Readiness**: The child demonstrates foundational skills and behaviors that prepare him or her for instruction based on kindergarten standards.
- **Approaching Readiness**: The child demonstrates some foundational skills and behaviors that prepare him or her for instruction based on kindergarten standards.
- **Emerging Readiness**: The child demonstrates minimal foundational skills and behaviors that prepare him or her for instruction based on kindergarten standards.

These PLDs are critical to establishing a common understanding of readiness and for supporting the standard setting activities that determine the cut scores for each of these levels. The process of standard setting establishes the aforementioned performance levels by setting two cut scores on the overall KRA scale.

A well-established standard setting procedure, known as Bookmark, was used for the KRA (Mitzel, Lewis, Patz, & Green, 2001). With this method, panelists review an ordered item booklet in which the content of the assessment is presented in the order of difficulty, from easiest to most difficult, based on how students actually performed on the items. The ordered item booklet allows panelists to set a cut score that represents the content that students should know or be able to do in order for them to be at a particular performance level.

The original standard setting meeting was held on February 18–20, 2015, and the panelists were educators from Maryland and Ohio. The ordered item booklet was based on the reduced number of items reflected in the KRA 1.5 Blueprint, and utilized the item data collected during the fall 2014 administration. The items were ordered using the response probability of 0.67 (i.e., RP67), which means that a student whose performance is at the cut score has a 2/3 chance of correctly answering the item at the cut score.

The fall 2015 administration of the KRA was the second complete census administration; however, it was the first administration that included only 50 items. (The fall 2014 administration included 63 items.) Also, the fall 2015 administration included numerous improvements to the overall administration, including enhancements to the technology system and professional development. Because this was the first administration of the 50-item version of the assessment with the enhancements and improvements to the administration process, a standard setting validation was conducted to ensure that the cut scores from the original standard setting were still appropriate. The standard setting validation meeting took place February 4–5, 2016, in Columbus, Ohio.

3.4.2 Method

A similar procedure (i.e., Bookmark method) was used for the standard setting validation. The goal of the validation process was for panelists to review the ordered item booklet and cut scores that were established during the original standard setting in order to determine if the cut scores needed to be adjusted. The panelists for this process were a group of 13 educators from Maryland and Ohio who had not participated in the original standard setting.

At the start of the meeting, the panelists were provided an overview of standard setting and its purpose, and they were trained on the process of placing bookmarks within the ordered item booklet. In this case, the training explained how each panelist would place two bookmarks within the ordered item booklet (i.e., cut scores) in order to establish the three performance levels. The first bookmark would be used to identify the items that separate students from the emerging to approaching readiness levels, and the second bookmark would be used to identify the items that separate students who are approaching readiness from those who are demonstrating readiness.

After the overview and training, the panelists spent the remainder of the first day reviewing the ordered item booklet in detail in order to familiarize themselves with the content. Specifically, the panelists used an aligned item map to take notes and document the accumulation of skills and behaviors that a student needed to possess in order to correctly answer the items that appeared later in the booklet (later items are more difficult than those that precede them). The panelists were split into two groups for this review.

Following the thorough review of the ordered item booklet, the whole group of panelists discussed the skills and knowledge, as described by the ordered item booklet and the previously established cut scores, that a student who is just entering a particular performance level is expected to master. The key distinction between the performance levels focused on the degree of remediation or support that a student required. Students in the approaching readiness level were described as those who could often demonstrate skills and behaviors with some adult assistance or support. Students in the demonstrating readiness were described as those who could demonstrate skills and behaviors independently and fluently, requiring little to no remediation. These students were characterized as “target students” for the approaching readiness and demonstrating readiness levels.

Upon conclusion of the whole-group discussion about the target students, the panelists independently set their bookmarks for round one. Panelists submitted their recommendations for the cut scores, and the median of the bookmark placements was determined for each table and for the overall group. (The median is selected as the best indicator of the group because it is not sensitive to extreme values, as is the mean.) The results of the first round were presented, and then the panelists engaged in discussions about the outcomes at their respective tables.

Once the table discussions were completed, the panelists set their second set of bookmarks (Round 2). They were encouraged to consider the group discussion when making their second selections, but panelists still submitted their Round 2 bookmarks independently. The medians and ranges of the Round 2 bookmarks were calculated and shared with the group. For this round, the impact data were shared in addition to the medians and ranges by table and overall group. Impact data indicated the proportion of

students in the combined population of Maryland and Ohio who would fall into each category based on their recommended cuts from Round 2.

The panelists then engaged in a whole-group discussion of the Round 2 results and their associated impact data. During this discussion, the panelists were encouraged to consider the relationship to the original cut scores in order to determine if the original cut scores needed to be adjusted. Upon completion of the whole-group discussion, the panelists independently set their final recommendations for the cut scores.

3.4.3 Results

The results of the final round aligned with the originally established cut scores (i.e., the median cut scores for approaching readiness and demonstrating readiness corresponded to the original cut scores). Table 3.3.3 includes a summary of the median, minimum, and maximum cut scores for all three rounds.

Table 3.3.3—Summary of Cut Scores for All Standard Setting Validation Rounds

	Approaching Readiness			Demonstrating Readiness		
	Minimum	**Median**	**Maximum**	**Minimum**	**Median**	**Maximum**
Round 1	253	257	260	265	270	273
Round 2	257	257	260	267	270	273
Round 3	257	257	258	269	270	270

3.4.4 Third Grade Reading Guarantee in Ohio

The Ohio panelists were also guided through the same process described above in order to review the cut score for the Language and Literacy domain, which supports the Third Grade Reading Guarantee within the state. The Ohio panelists participated in two rounds, utilizing an ordered item booklet that consisted of only the Language and Literacy items. The results also supported the retention of the original cut score for the Third Grade Reading Guarantee. Table 3.3.4 includes a summary of the median, minimum, and maximum cut scores for both rounds.

Table 3.3.4—Summary of Cut Scores for All Rounds for the Third Grade Reading Guarantee

	Third Grade Reading Guarantee—Language and Literacy Domain		
	Minimum	**Median**	**Maximum**
Round 1	260	263	266
Round 2	263	263	263

4 Technology Support

4.1 Overview of Ready for Kindergarten Online System

Technical development of the Ready for Kindergarten online system was led by the Johns Hopkins University Center for Technology in Education (JHU CTE).

The Ready for Kindergarten online system comprises two key components: (1) the Ready for Kindergarten website (the primary teacher interface) and (2) the Ready for Kindergarten App (for delivering a subset of the KRA items directly to students, using child-friendly technologies). The Ready for Kindergarten online system supports:

- administration and scoring of the KRA, including a subset of the KRA items using child-friendly, touch-screen technologies;
- import and export of data to and from state longitudinal data systems;
- reports summarizing student-level results;
- reports to monitor completion of the KRA by key personnel at the local, district, and state levels;
- back-end data management of enrollment and school information;
- management of the assessment content and supporting materials; and
- delivery and support of professional development.

4.2 Improvements to Ready for Kindergarten Online System

Several improvements and enhancements were made to the Ready for Kindergarten online system between the fall 2014 administration and the fall 2015 administration of the KRA. The enhanced version (i.e., Version 1.5) of the Ready for Kindergarten online system was launched on August 1, 2015, to kindergarten programs across Maryland and Ohio.

The improvements and enhancements to the Ready for Kindergarten online system included:

- eight new items within the Ready for Kindergarten App, bringing the total to 17 items and a tutorial/practice item;
- printable score sheets, including student names, for recording item scores;
- default for item scores was changed to "*Needs to be Administered*" in order to assist teachers with keeping track of student data in order to reduce the number of blank item scores;
- ability to enter item scores for the entire assessment on one screen (scores are auto-saved each time a score is entered to assist data entry and reduce teacher burden);
- dashboard access to student results for immediate use by teachers;
- capability to assign more than one teacher to a student for data entry;
- Individual Student Reports (ISRs), available directly within the system upon completion of the KRA administration (in November); and
- Individual Student Reports available in multiple languages, including English, Spanish, Chinese, and French.

5 Professional Development

5.1 Overview of Professional Development Approach

Professional development is one aspect of overall implementation that requires intentional design and customized delivery of information around the assessment and technology systems. Implementation includes careful attention to individual state needs and support for all stakeholders interacting directly with the KRA. In order for JHU CTE to effectively implement the professional development, the following strategies were employed.

- Learn about each state's unique needs, policies, and processes so that the professional development process is effectively implemented to scale.
- Implement a multilevel evaluation strategy to promote training and assessment implementation fidelity that includes simulation technology, surveys, and fidelity checklists.
- Offer professional development through a variety of formats to engage relevant audiences.
- Use online communities to support interaction among audience members and to promote resource sharing.
- Collaborate with the assessment and technology teams to ensure that the professional development process effectively supports the system.
- Provide ongoing consultation to states, as well as support to local leaders and trainers through online FAQs, communication, and meetings to address ongoing implementation questions and challenges as they arise.

5.2 Improvements to Professional Development

Several improvements and enhancements were made to the professional development modules and resources for the fall 2015 administration of the KRA. The improvements and enhancements to professional development included:

- updated support information about the four domains and 50 KRA items, including tips for administering the KRA effectively;
- a new Observation Planning Guide, aligned to the Teacher Administration Manual;
- new resources for integrating the KRA throughout a typical classroom day;
- updated Universally Designed Allowances documentation, supporting activities, and resources;
- an updated Guidelines for Allowable Supports document, corresponding activities, and resources for supporting children who require support;
- a new set of videos for practicing and scoring observational-rubric items;
- additional domain-specific resources for supporting the essential skills, knowledge, and behaviors within the Common Language Standards;
- guidelines for teachers working with proctors who helped to administer KRA App items to the children in their classrooms;
- an updated Teacher Administration Manual for the Blind and Visually Impaired, including foam and tactile manipulatives.

References

Carmines, E. G., & Zeller, R. A. (1979). *Reliability and validity assessment.* Newbury Park, CA: Sage.

Mitzel, H. C., Lewis, D. M., Patz, R. J., & Green, D. R. (2001). The Bookmark procedure: Psychological perspectives. In Cizek, G. J. (Ed.), *Setting Performance Standards: Concepts, Methods and Perspectives* (pp. 249–281). Mahwah, NJ: Lawrence Earlbaum Associates.

Appendices

Appendix A: Project Members by Organization

Maryland State Department of Education (MSDE)

Rolf Grafwallner, Ph.D.	Assistant State Superintendent, Division of Early Childhood Development
Judy Walker	Early Learning Branch Chief, Division of Early Childhood Development
Candy Miller	Early Learning Program and Assessment Specialist, Division of Early Childhood Development
Robert Wagner	Early Learning Education Program Specialist, Division of Early Childhood Development
Marcella Franczkowski	Assistant State Superintendent, Division of Special Education and Early Intervention Services
Nancy Vorobey	Section Chief, Division of Special Education and Early Intervention Services

Ohio Department of Education (ODE)

Stephanie Siddens, Ph.D.	Senior Executive Director, Center for Curriculum & Assessment
Wendy Grove, Ph.D.	Director, Early Learning & School Readiness
Lauren Monowar-Jones, Ph.D.	Assistant Director, Early Learning & School Readiness

Johns Hopkins University Center for Technology in Education (JHU CTE)

Jacqueline Nunn, Ed.D.	Director
Christopher Sessums, Ph.D.	Director of Research and Evaluation
Dave Peloff	Senior Program Director of Technology
Linda Carling	Program Director of Online Learning
Kristen Thompson	Technology Program Coordinator
Angela Vann	Implementation Specialist for Professional Development

WestEd—Standards, Assessment, and Accountability Services (SAAS) Program

Andrew Latham, Ph.D.	Director
Joanne Jensen, Ph.D.	Director of Assessment Client Relations
Matthew Brunetti	Project Manager
Beverly Nedrow	Senior Content Specialist
Amanda Otte	Content Specialist
Wenson Fung, Ph.D.	Research Assistant

Psychometric Consultant

Karla Egan, Ph.D.	Psychometrician

Technical Advisory Committee
Facilitated by the Council of Chief State School Officers

Jerry West, Ph.D. (Chair)	Senior Fellow, Mathematica Policy Research
Clancy Blair, Ph.D.	Professor of Applied Psychology, Steinhardt School of Culture Education and Human Development, New York University
Linda Espinosa, Ph.D.	Co-Principal Investigator, Getting on Track for School Success, University of Chicago, National Opinion Research Center
Daryl B. Greenfield, Ph.D.	Professor of Psychology & Pediatrics, Department of Psychology, University of Miami
Ilonca Hardy, Ph.D.	Department of Education, Goethe University
Kathleen Hebbeler, Ph.D.	Program Manager, Community Services and Strategies, SRI International
Nancy Jordan, Ed.D.	Professor of Education, University of Delaware
Laura Justice, Ph.D.	Executive Director, The Schoenbaum Family Center**,** Ohio State University
Sharon Lynn Kagan, Ed.D.	Professor of Early Childhood and Family Policy**,** Co-Director of the National Center for Children and Families, Teachers College, Columbia University
Robert Lissitz, Ph.D.	Professor of Education, University of Maryland
Catherine Scott-Little, Ph.D.	Associate Professor, Human Development and Family Studies**,** School of Health and Human Sciences, The University of North Carolina at Greensboro
C. Cybele Raver, Ph.D.	Professor of Applied Psychology and Developmental Psychology**,** Steinhardt School of Culture Education and Human Development, New York University
Ross Thompson, Ph.D.	Distinguished Professor of Psychology**,** University of California, Davis

Appendix B: Item-Level Classical Statistics

Item Code	Type	Max	M	SD	Difficulty (*p*-value)	Disc. (Item-Rest)	Number of Scored Responses	Students at Score Point (%)				N = 188,118	
								0	1	2	3	Not Scorable (%)	Missing (%)
MA.1.1.A_A101	PT	3	2.31	0.99	0.77	0.57	180363	6.36	18.90	12.61	62.13	0.68	3.44
MA.1.1.C_A104	PT	3	2.10	0.88	0.70	0.53	180386	6.15	15.89	40.13	37.83	0.62	3.49
MA.1.1.D_A121	PT	2	1.84	0.45	0.92	0.39	180362	3.58	8.86	87.56		0.63	3.49
MA.1.1.F_A115	PT	1	0.83	0.37	0.83	0.36	180367	16.76	83.24			0.63	3.49
MA.1.1.G_A117	PT	3	2.15	1.06	0.72	0.45	180580	11.89	14.16	20.85	53.10	0.57	3.44
MA.2.1.B_A138	PT	1	0.46	0.50	0.46	0.33	180587	54.14	45.86			0.57	3.43
MA.3.1.B_A123	PT	2	1.45	0.87	0.72	0.27	180720	25.67	3.82	70.51		0.54	3.40
MA.3.1.D_A143	SR	1	0.82	0.39	0.82	0.29	180571	18.20	81.80			0.57	3.44
MA.3.1.D_A147	SR	1	0.53	0.50	0.53	0.44	180537	47.21	52.79			0.58	3.45
MA.3.1.D_A149	SR	1	0.54	0.50	0.54	0.26	180558	46.17	53.83			0.58	3.44
MA.3.2.A_A152	PT	2	1.57	0.65	0.79	0.42	180687	9.09	24.77	66.14		0.55	3.40
MA.3.2.B_A174	PT	1	0.44	0.50	0.44	0.36	180608	56.25	43.75			0.59	3.40
MA.4.1.A_A177	PT	2	1.36	0.65	0.68	0.23	180782	9.64	44.53	45.83		0.49	3.40
MA.4.1.B_A191	PT	2	1.51	0.65	0.75	0.49	180206	8.45	32.42	59.13		0.71	3.50
LL.1.1.A_H101	SR	1	0.91	0.29	0.91	0.23	180584	8.99	91.01			0.59	3.42
LL.1.1.B_H104	SR	1	0.93	0.26	0.93	0.26	180552	7.44	92.56			0.60	3.42
LL.1.1.C_H106R	PT	3	1.80	1.05	0.60	0.47	180238	15.19	21.59	31.61	31.60	0.74	3.45
LL.1.2.A_A163	SR	1	0.61	0.49	0.61	0.28	180603	39.29	60.71			0.60	3.39
LL.1.2.B_A164	PT	3	1.71	1.05	0.57	0.29	180090	18.87	17.68	36.99	26.45	0.79	3.48
LL.1.2.D_A127	SR	1	0.65	0.48	0.65	0.42	180604	34.60	65.40			0.61	3.39
LL.1.2.D_A180	SR	1	0.69	0.46	0.69	0.27	180585	31.19	68.81			0.60	3.40
LL.1.3.B_A130	PT	2	1.20	0.80	0.60	0.52	180127	24.47	31.31	44.23		0.74	3.51
LL.1.3.C_A132	PT	3	2.09	1.07	0.70	0.55	180359	12.48	15.46	22.80	49.26	0.64	3.48
LL.3.1.A_A134	PT	2	1.73	0.56	0.86	0.55	180564	5.79	15.54	78.67		0.55	3.46
LL.3.1.B_A136	PT	2	1.82	0.47	0.91	0.48	180492	3.83	10.51	85.66		0.57	3.48
LL.4.1.A_A155	PT	3	2.72	0.59	0.91	0.46	180278	1.28	3.35	17.74	77.63	0.70	3.46
LL.4.1.A_A195	PT	3	2.52	0.69	0.84	0.44	180159	1.88	5.46	31.03	61.64	0.77	3.46
LL.4.1.D_A160	PT	3	2.56	0.75	0.85	0.47	180705	2.13	9.14	18.89	69.84	0.56	3.38
LL.4.2.B_H103R	SR	1	0.65	0.48	0.65	0.28	180508	34.66	65.34			0.62	3.43

Note: Percentages for Not Scorable and Missing are based on the total population (N = 188,118). Percentages for Students at Score Point are based on the number of scored responses.

Item Code	Type	Max	M	SD	Difficulty (*p*-value)	Disc. (Item-Rest)	Number of Scored Responses	Students at Score Point (%)				N = 188,118	
								0	1	2	3	Not Scorable (%)	Missing (%)
LL.2.1.A_OR_19	OR	2	1.55	0.64	0.78	0.61	180179	7.95	28.82	63.23		0.66	3.56
LL.2.1.B_OR_20	OR	2	1.56	0.65	0.78	0.65	180135	8.76	26.50	64.74		0.68	3.56
SF.1.1.B_OR_35	OR	2	1.61	0.61	0.80	0.59	180221	6.58	26.35	67.08		0.60	3.60
SF.1.2.B_OR_42	OR	2	1.51	0.65	0.75	0.61	180180	8.39	32.43	59.18		0.61	3.61
SF.1.2.C_OR_43	OR	2	1.66	0.58	0.83	0.60	180219	5.44	23.25	71.31		0.58	3.62
SF.2.1.C_OR_49	OR	2	1.64	0.59	0.82	0.45	180538	5.79	24.05	70.16		0.45	3.58
SF.2.2.B_OR_55	OR	2	1.30	0.70	0.65	0.58	180502	14.30	41.29	44.40		0.46	3.59
SF.2.3.A_OR_58	OR	2	1.37	0.70	0.69	0.68	180444	12.72	37.11	50.17		0.49	3.59
SF.2.3.C_OR_12	OR	2	1.31	0.72	0.65	0.66	180369	15.23	38.70	46.07		0.50	3.62
SF.2.5.A_OR_25	OR	2	1.57	0.61	0.78	0.59	180141	6.06	31.13	62.81		0.50	3.74
SF.2.5.A_OR_64	OR	2	1.45	0.65	0.73	0.63	180283	8.81	36.92	54.27		0.55	3.61
SF.2.6.B_OR_68	OR	2	1.63	0.58	0.81	0.54	180084	4.90	27.43	67.67		0.61	3.66
SF.2.6.D_OR_69	OR	2	1.72	0.53	0.86	0.47	180327	4.04	20.10	75.85		0.44	3.70
SF.3.1.B_OR_32	OR	2	1.61	0.61	0.81	0.66	179753	6.92	24.74	68.34		0.74	3.71
PD.1.1.A_OR_01	OR	2	1.76	0.51	0.88	0.46	180297	3.81	16.36	79.83		0.44	3.72
PD.1.1.B_OR_05	OR	2	1.73	0.53	0.87	0.45	180105	4.48	17.61	77.91		0.50	3.76
PD.1.2.B_OR_06	OR	2	1.61	0.62	0.80	0.55	180271	7.08	25.04	67.88		0.49	3.68
PD.1.2.C_OR_08	OR	2	1.67	0.57	0.84	0.51	180457	5.42	21.72	72.86		0.39	3.68
PD.2.1.A_OR_10	OR	2	1.70	0.54	0.85	0.46	180323	4.31	21.55	74.14		0.43	3.71
PD.2.1.B_OR_09	OR	2	1.58	0.62	0.79	0.62	179733	7.35	26.81	65.83		0.70	3.75
PD.2.2.A_OR_15	OR	2	1.77	0.47	0.88	0.49	180327	2.27	18.68	79.05		0.45	3.69

Note: Percentages for Not Scorable and Missing are based on the total population (N = 188,118). Percentages for Students at Score Point are based on the number of scored responses.

Appendix C: Item-Level Statistics by KRA App Use

Mathematics Domain

No App

Item Code	Type	Max	M	SD	Difficulty (*p*-value)	Disc. (Item-Rest)	Students at Score Point (%)					N = 101,898	
							Number of Scored Responses	0	1	2	3	Not Scorable (%)	Missing (%)
MA.1.1.G_A117	PT	3	1.95	1.17	0.65	0.49	94400	19.04	14.41	19.34	47.21	1.02	6.34
MA.2.1.B_A138	PT	1	0.46	0.50	0.46	0.40	94401	53.66	46.34			1.02	6.33
MA.3.1.B_A123	PT	2	1.63	0.73	0.82	0.34	94535	15.36	5.78	78.86		0.96	6.27
MA.3.1.D_A143	SR	1	0.88	0.33	0.88	0.36	94393	12.09	87.91			1.01	6.35
MA.3.1.D_A147	SR	1	0.56	0.50	0.56	0.51	94367	43.65	56.35			1.03	6.36
MA.3.1.D_A149	SR	1	0.55	0.50	0.55	0.38	94380	45.25	54.75			1.04	6.34
MA.3.2.A_A152	PT	2	1.62	0.62	0.81	0.44	94500	7.53	23.26	69.21		0.99	6.27
MA.3.2.B_A174	PT	1	0.50	0.50	0.50	0.42	94420	50.50	49.50			1.07	6.27
MA.4.1.A_A177	PT	2	1.62	0.55	0.81	0.28	94586	3.26	31.30	65.43		0.89	6.28

App **N = 84,016**

MA.1.1.G_A117	PT	3	2.38	0.87	0.79	0.23	84016	3.94	13.86	22.53	59.67	---	---
MA.2.1.B_A138	PT	1	0.45	0.50	0.45	0.31	84016	54.65	45.35			---	---
MA.3.1.B_A123	PT	2	1.24	0.96	0.62	0.25	84016	37.14	1.57	61.29		---	---
MA.3.1.D_A143	SR	1	0.75	0.43	0.75	0.23	84016	24.93	75.07			---	---
MA.3.1.D_A147	SR	1	0.49	0.50	0.49	0.40	84016	51.13	48.87			---	---
MA.3.1.D_A149	SR	1	0.53	0.50	0.53	0.20	84016	47.22	52.78			---	---
MA.3.2.A_A152	PT	2	1.52	0.68	0.76	0.32	84016	10.79	26.39	62.81		---	---
MA.3.2.B_A174	PT	1	0.37	0.48	0.37	0.37	84016	62.68	37.32			---	---
MA.4.1.A_A177	PT	2	1.07	0.63	0.53	0.28	84016	16.75	59.60	23.65		---	---

Note: Percentages for Not Scorable and Missing are based on the total population (No App: N = 101,898; App: N = 84,016).
Percentages for Students at Score Point are based on the number of scored responses.
App items did not have Not Scorable or Missing.

Language and Literacy Domain

No App

Item Code	Type	Max	M	SD	Difficulty (*p*-value)	Disc. (Item-Rest)	Students at Score Point (%)					N = 105,845	
							Number of Scored Responses	0	1	2	3	Not Scorable (%)	Missing (%)
LL.1.1.A_H101	SR	1	0.93	0.26	0.93	0.30	98412	7.19	92.81			0.99	6.03
LL.1.1.B_H104	SR	1	0.94	0.23	0.94	0.34	98390	5.75	94.25			1.01	6.03
LL.1.1.C_H106R	PT	3	1.75	1.07	0.58	0.44	98199	16.97	21.33	31.07	30.62	1.18	6.04
LL.1.2.A_A163	SR	1	0.62	0.49	0.62	0.31	98357	38.13	61.87			1.05	6.03
LL.1.2.D_A127	SR	1	0.66	0.47	0.66	0.42	98363	33.93	66.07			1.05	6.01
LL.1.2.D_A180	SR	1	0.68	0.47	0.68	0.33	98339	31.63	68.37			1.05	6.04
LL.4.1.D_A160	PT	3	2.63	0.74	0.88	0.46	98451	2.62	7.52	14.09	75.78	0.97	6.01
LL.4.2.B_H103R	SR	1	0.69	0.46	0.69	0.35	98355	31.40	68.60			1.03	6.05

App **N = 68,041**

LL.1.1.A_H101	SR	1	0.89	0.32	0.89	0.23	68041	11.34	88.66			---	---
LL.1.1.B_H104	SR	1	0.91	0.29	0.91	0.24	68041	9.46	90.54			---	---
LL.1.1.C_H106R	PT	3	1.85	1.02	0.62	0.38	68041	13.07	21.80	32.19	32.94	---	---
LL.1.2.A_A163	SR	1	0.59	0.49	0.59	0.20	68041	40.75	59.25			---	---
LL.1.2.D_A127	SR	1	0.65	0.48	0.65	0.38	68041	35.05	64.95			---	---
LL.1.2.D_A180	SR	1	0.69	0.46	0.69	0.23	68041	30.67	69.33			---	---
LL.4.1.D_A160	PT	3	2.50	0.74	0.83	0.40	68041	1.48	10.72	24.60	63.19	---	---
LL.4.2.B_H103R	SR	1	0.61	0.49	0.61	0.27	68041	39.31	60.69			---	---

Note: Percentages for Not Scorable and Missing are based on the total population (No App: N = 105,845; App: N = 68,041).
Percentages for Students at Score Point are based on the number of scored responses.
App items did not have Not Scorable or Missing.

Appendix D: Distributions of Scaled Scores

Table D.1 – Distribution of KRA Overall Scaled Scores

Scaled Score Range	Percent	Cumulative Percent
202-210	0.27	0.27
211-220	0.10	0.37
221-230	0.27	0.65
231-240	1.42	2.06
241-250	6.72	8.78
251-260	19.81	28.60
261-270	32.46	61.06
271-280	25.18	86.24
281-290	10.02	96.27
291-298	3.73	100.00
Mean		267.21
SD		13.12
N		178065

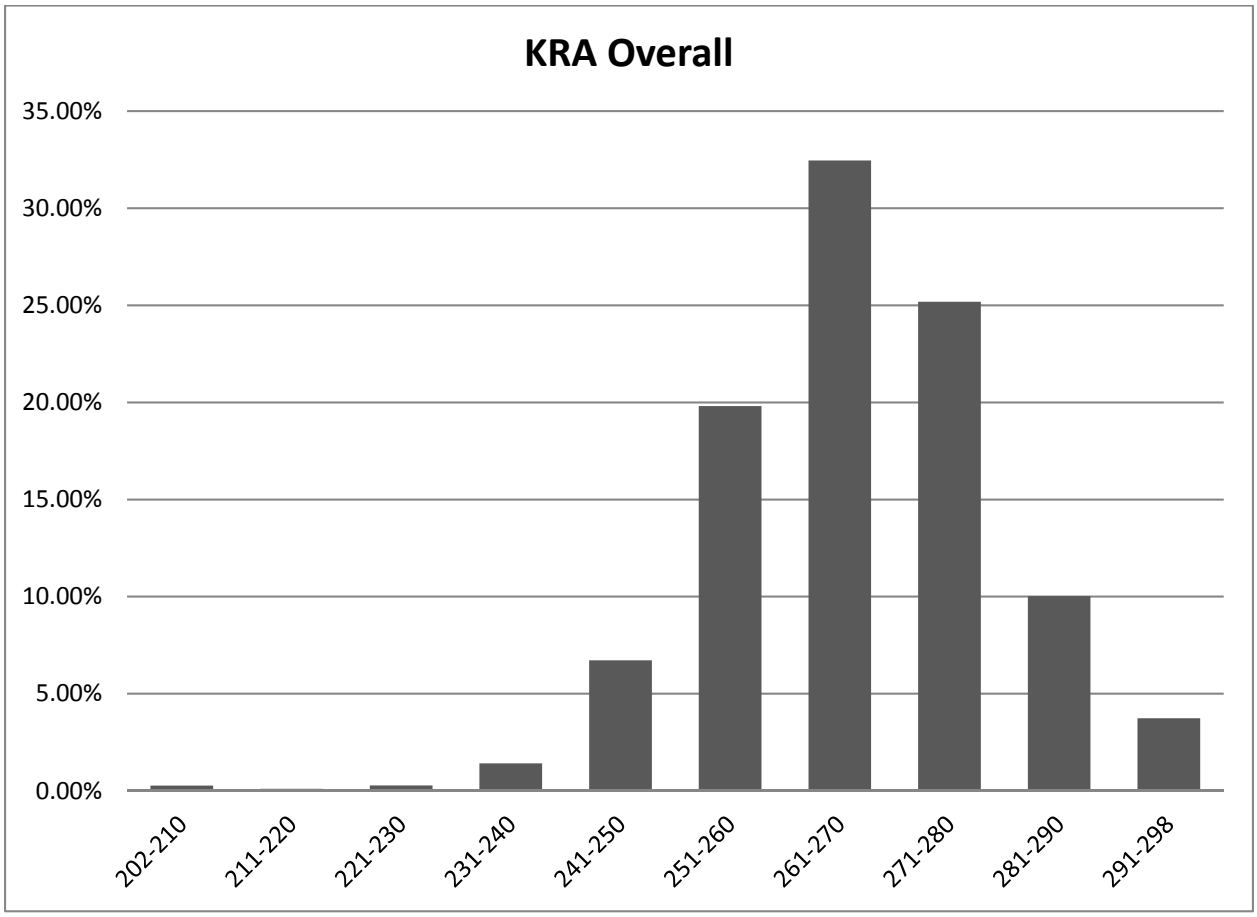


Table D.2 – Distribution of Language and Literacy Domain Scaled Scores

Scaled Score Range	Percent	Cumulative Percent
202-210	0.43	0.43
211-220	0.18	0.61
221-230	0.44	1.05
231-240	1.86	2.90
241-250	6.96	9.86
251-260	19.12	28.99
261-270	33.14	62.12
271-280	25.00	87.12
281-290	6.28	93.40
291-298	6.60	100.00
Mean		266.78
SD		13.83
N		181124

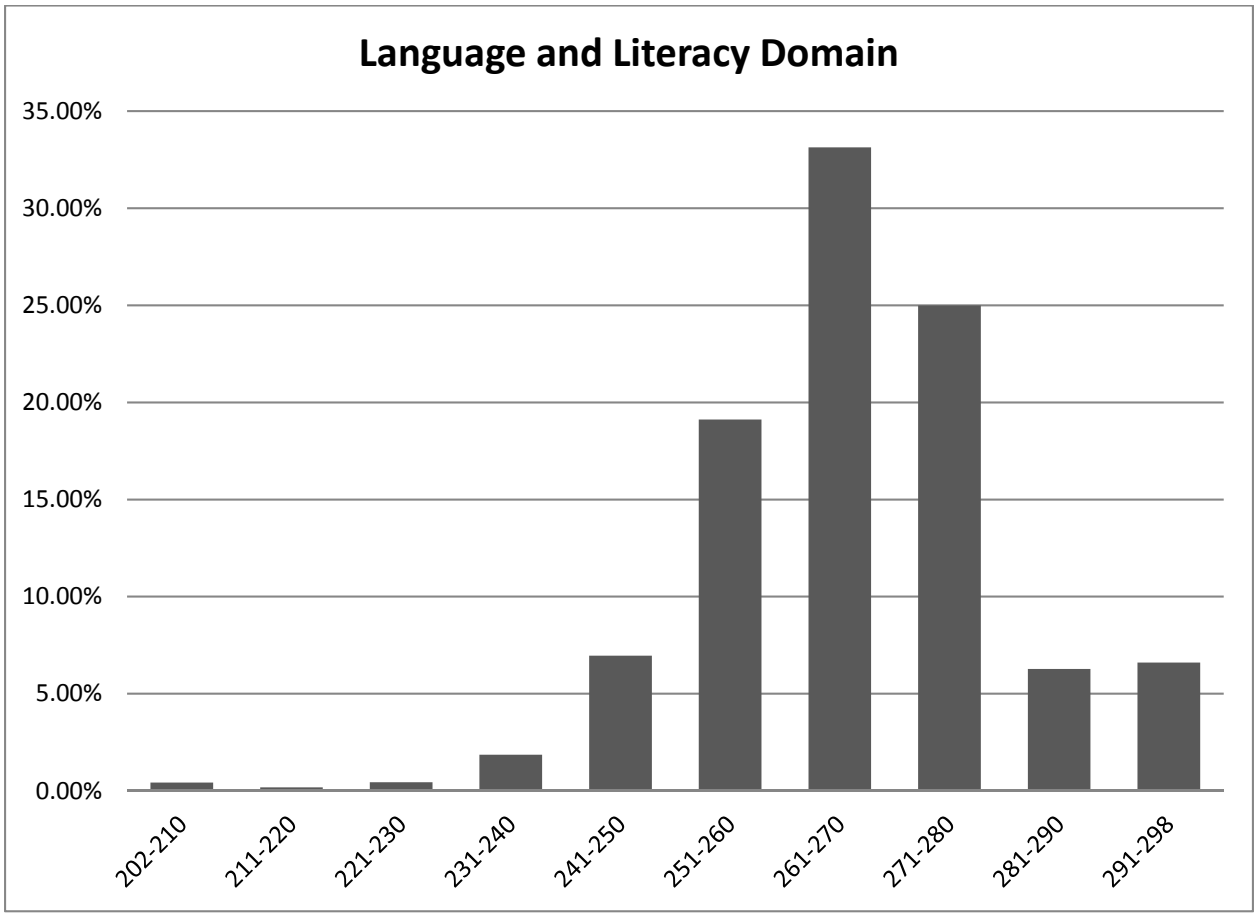


Table D.3 – Distribution of Mathematics Domain Scaled Scores

Scaled Score Range	Percent	Cumulative Percent
202-210	0.42	0.42
211-220	0.10	0.52
221-230	0.43	0.95
231-240	1.49	2.45
241-250	8.28	10.73
251-260	23.26	33.99
261-270	32.19	66.18
271-280	16.88	83.06
281-290	13.24	96.30
291-298	3.70	100.00
Mean		265.77
SD		14.08
N		182009

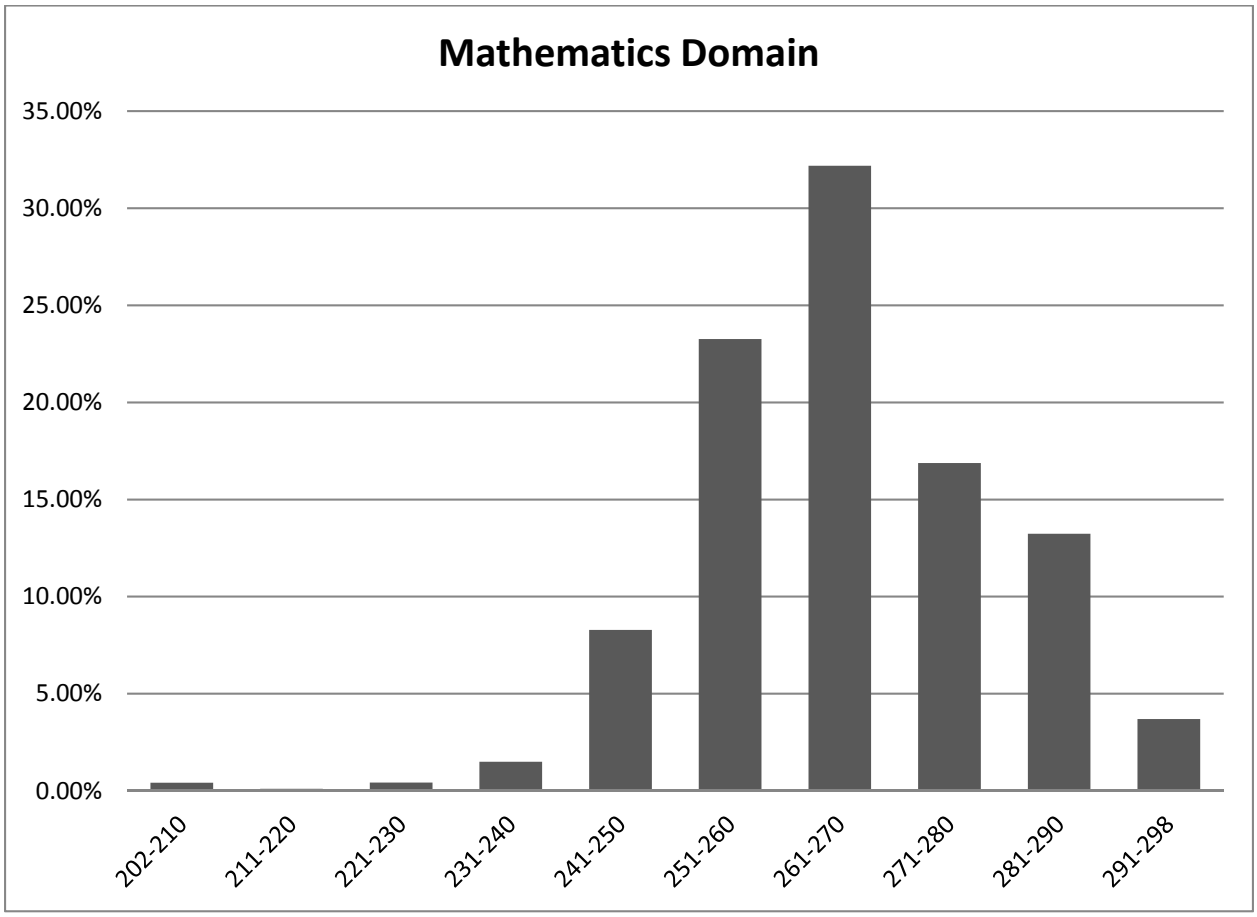
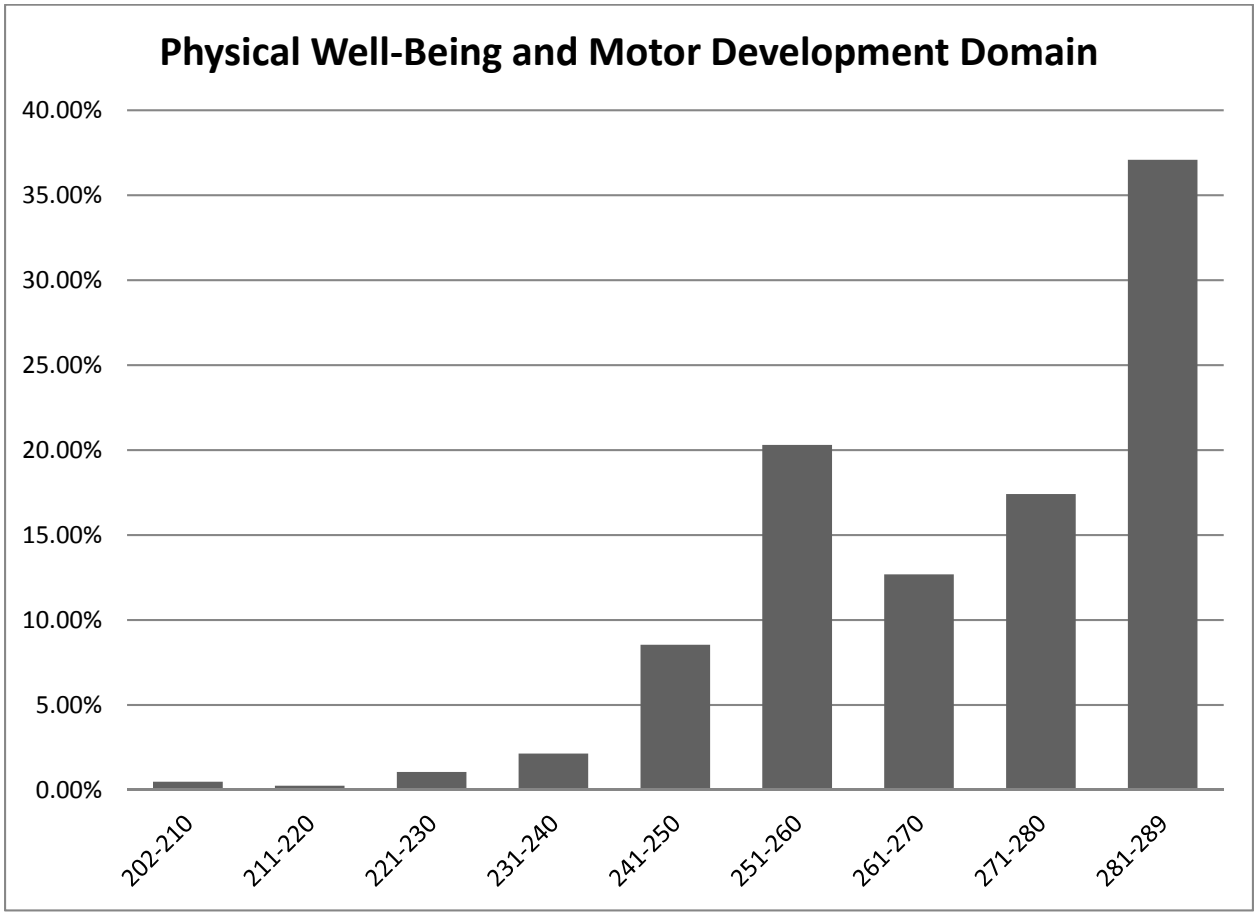


Table D.4 – Distribution of Physical Well-Being and Motor Development Domain Scaled Scores

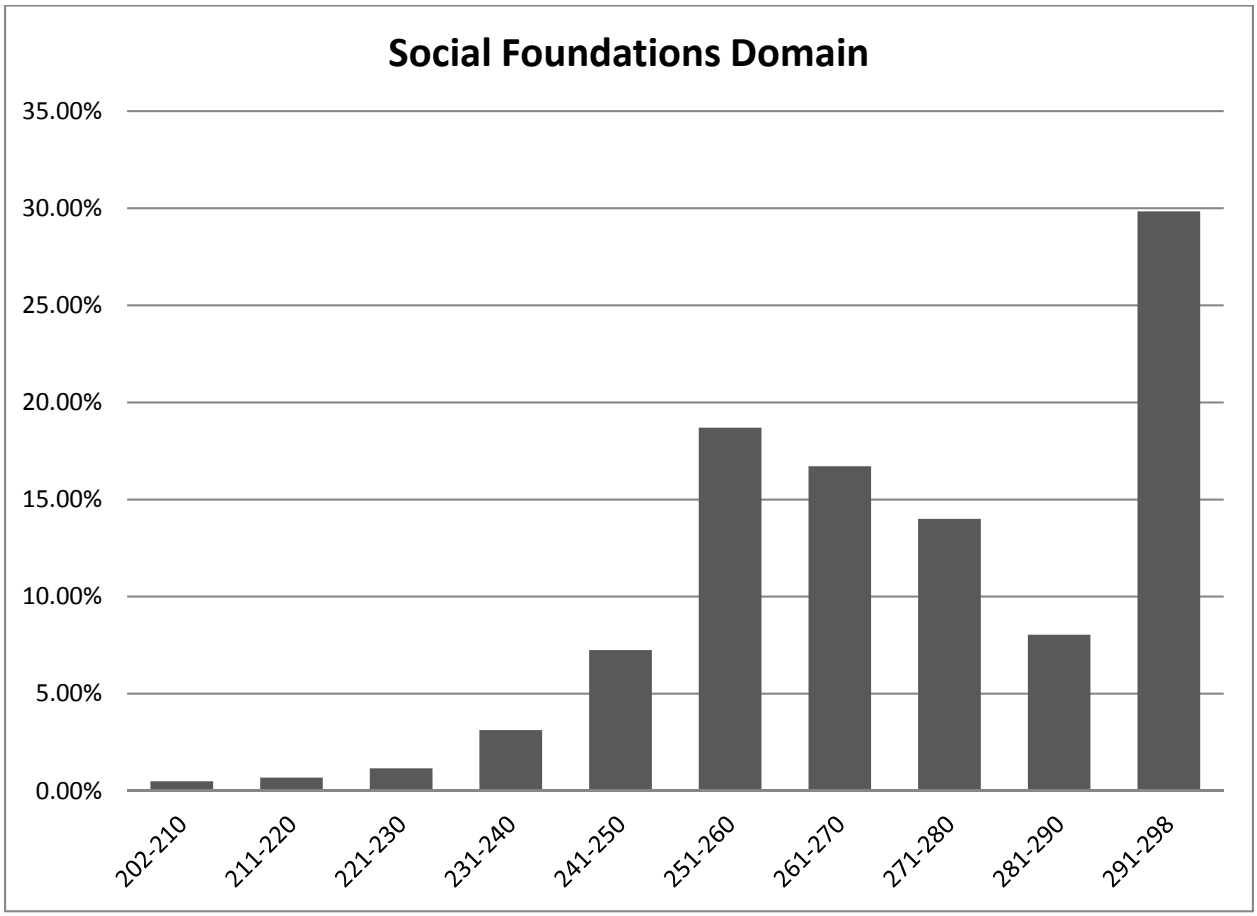
Scaled Score Range	Percent	Cumulative Percent
202-210	0.49	0.49
211-220	0.25	0.74
221-230	1.06	1.80
231-240	2.14	3.95
241-250	8.55	12.49
251-260	20.31	32.80
261-270	12.69	45.49
271-280	17.41	62.91
281-289	37.09	100.00
Mean		270.80
SD		17.65
N		180,688



Note: The large percentage of students in the top score range is likely due to the limitation of the scale, which is a result of a limited number of items and score points within this domain.

Table D.5 – Distribution of Social Foundations Domain Scaled Scores

Scaled Score Range	Percent	Cumulative Percent
202-210	0.49	0.49
211-220	0.68	1.17
221-230	1.16	2.33
231-240	3.13	5.45
241-250	7.24	12.70
251-260	18.70	31.40
261-270	16.72	48.12
271-280	14.00	62.12
281-290	8.04	70.16
291-298	29.84	100.00
Mean		272.23
SD		20.152
N		180,969



Note: The large percentage of students in the top score range is likely due to the truncation of the scale.


R 4 K
ready for
KINDERGARTEN
MARYLAND


ready for
KINDERGARTEN
OHIO


JOHNS HOPKINS
UNIVERSITY
School of Education
Center for
Technology in
Education


WestEd