TO:	Members of the State Board of Education
FROM:	Carey M. Wright, Ed.D., Interim State Superintendent of Schools
DATE:	February 27, 2024
SUBJECT:	Update on Kindergarten Readiness Assessment

Purpose

The purpose of this item is to present Kindergarten Readiness Assessment results for 2023-2024 school year.

Background

The current Kindergarten Readiness Assessment (KRA) measures the school readiness of incoming kindergartners. Kindergarten teachers administer the KRA to students during the first few months of school. The assessment measures students' skills, knowledge, and behaviors across four domains: language and literacy, mathematics, social foundations, and physical well-being and motor development. In school year 2023-2024, the KRA was administered to all kindergarteners as a census administration. The Maryland State Department of Education is tasked with administering a kindergarten readiness assessment to all incoming kindergarten students that is racially and culturally unbiased under the Blueprint for Maryland's Future and Education Code § 7-210.

Executive Summary

The presentation will provide state-level results of the KRA 2023-2024 administration, including student group disaggregations and comparisons to prior years. The presentation will also include an overview of a study and report to the AIB on the Evaluation of Racial, Cultural, or Linguistic Bias of the Kindergarten Readiness Assessment.

Action

No action is required; this information is for discussion only.

Attachments

Kindergarten Readiness Assessment Results.pdf

200 West Baltimore Street Baltimore, MD 21201 | 410-767-0100   Deaf and hard of hearing use Relay.

marylandpublicschools.org

Presentation Outline

1. About the Kindergarten Readiness Assessment (KRA)

2. State Results

3. Report to Accountability and Implementation Board (AIB)
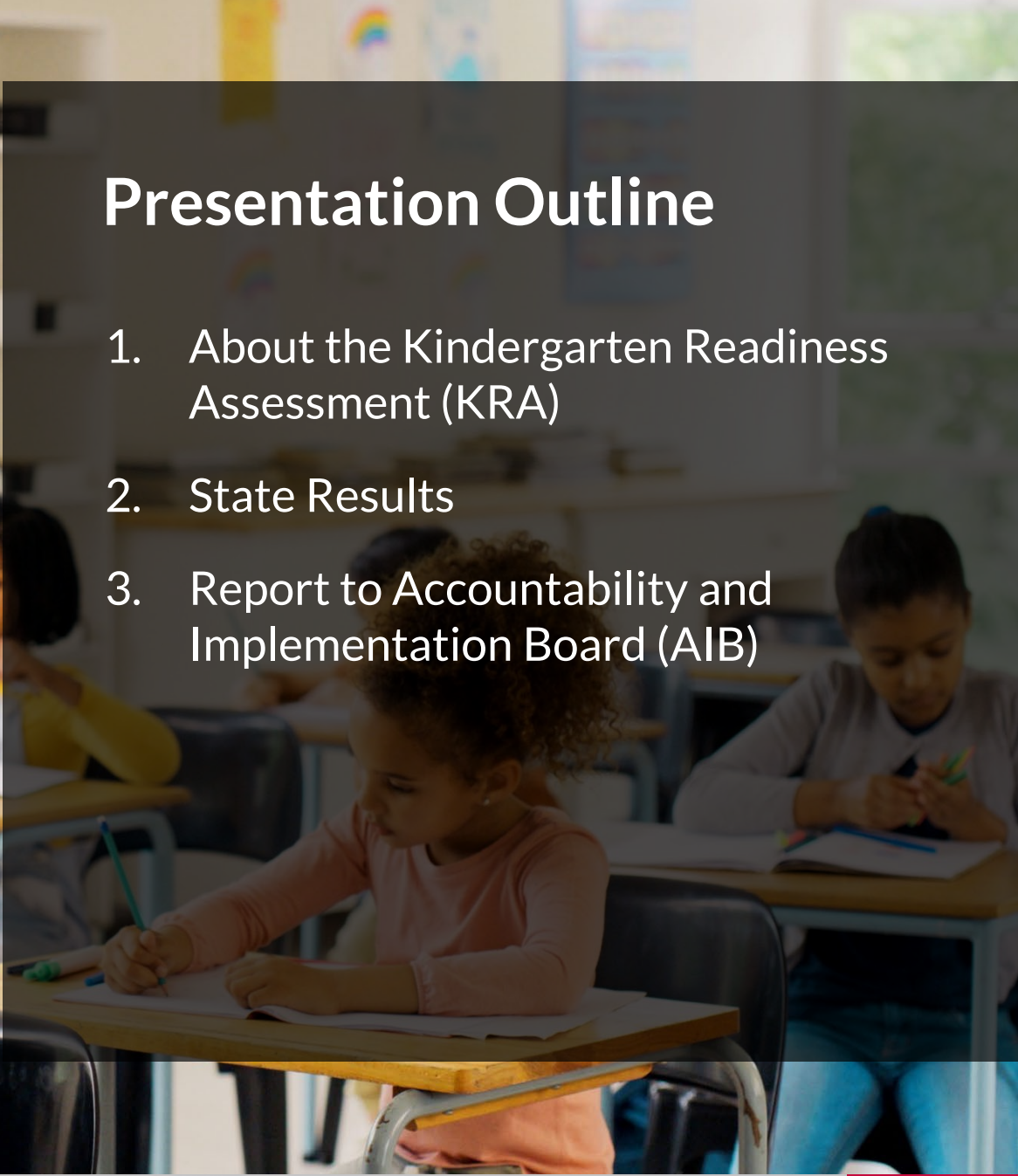
Kindergarten Readiness Assessment (KRA)

The current Kindergarten Readiness Assessment (KRA) is one component of Ready for Kindergarten (R4K): Maryland's Early Childhood Comprehensive Assessment System.

- The assessment uses observations of children's work and play, selected response items, and performance tasks.

- Specific kindergarten readiness skills are measured in four domains: language and literacy, social foundations, mathematics, and physical well-being and motor development.

- The current KRA is typically administered within the first two months of school.

- Kindergarteners are determined to be:

    - **Demonstrating Readiness** –consistently demonstrate the foundational skills and behaviors that enable a child to fully participate in the kindergarten curriculum.

    - **Approaching Readiness** –exhibit some of the foundational skills and behaviors that are needed to participate in the kindergarten curriculum.

    - **Emerging Readiness** –show minimal foundational skills and behaviors that prepare them to meet kindergarten expectations.

Maryland Statewide Kindergarten Readiness Over Time

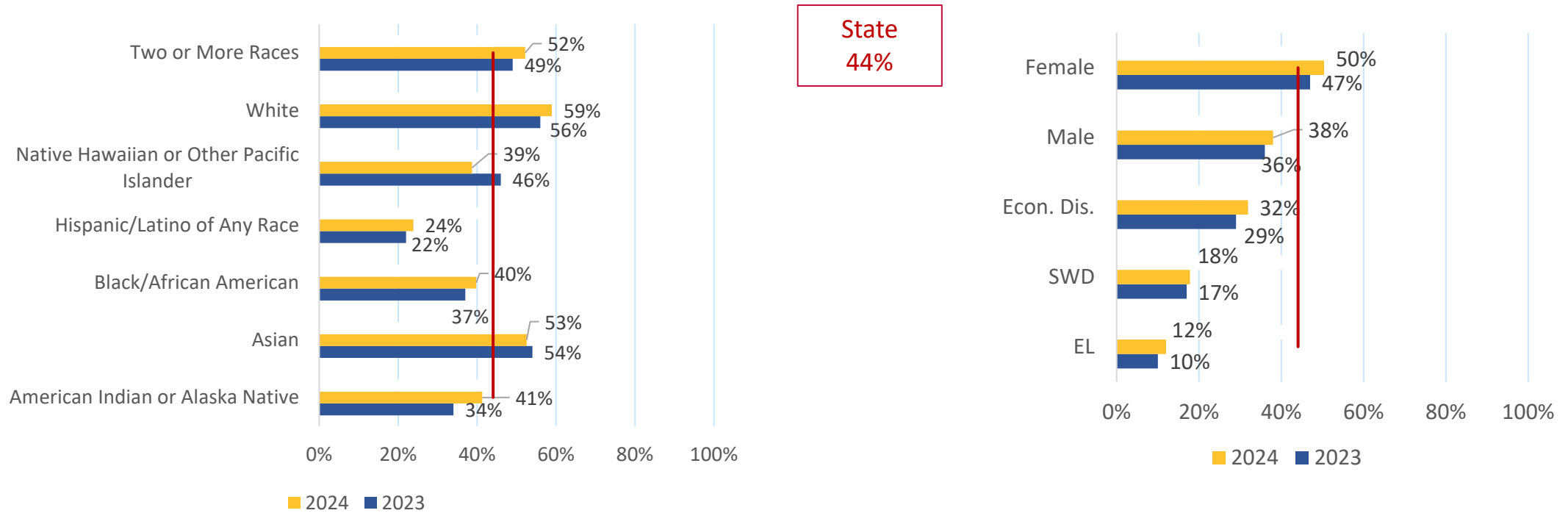**The percentage of students demonstrating readiness for kindergarten continued to recover from the pandemic, reaching 44% in the 2023-2024 school year.**
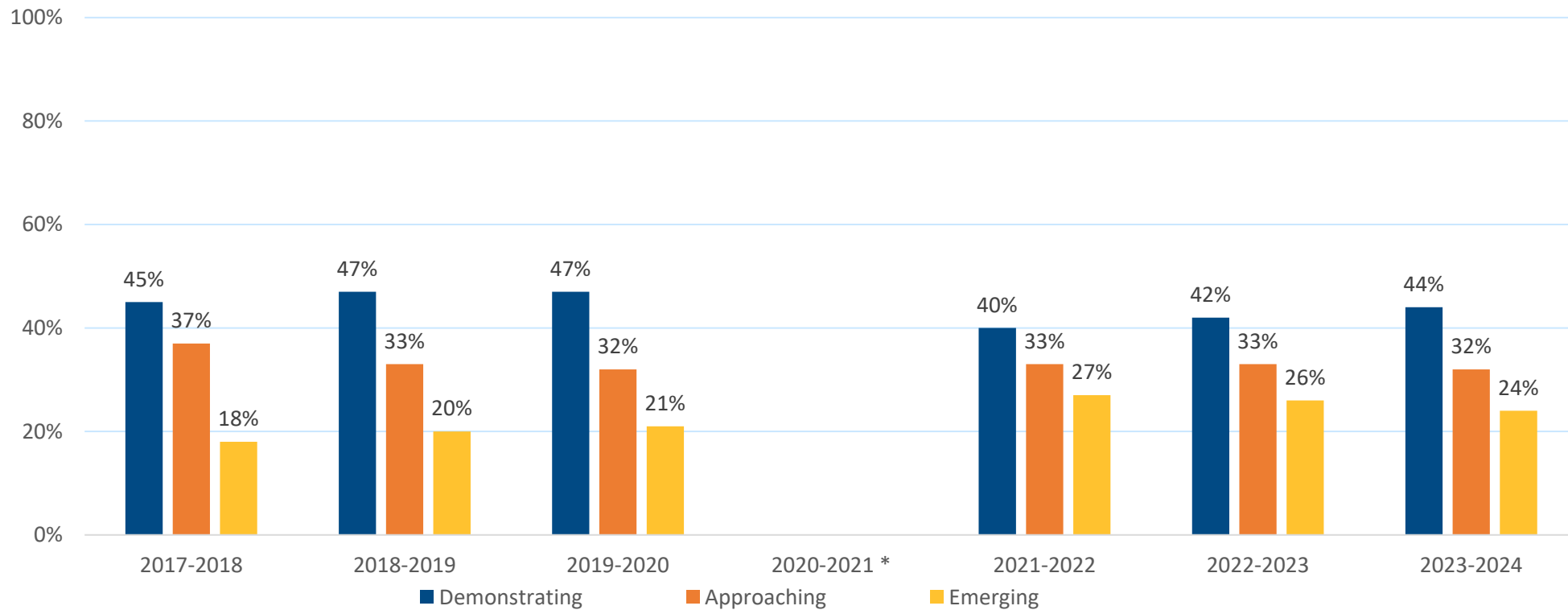
Note: *Due to the impact of the COVID-19 pandemic, Maryland did not administer the KRA in school year 2020-2021. Census administration of the KRA began in the 2022-2023 school year. Prior to that year, some LEAs were administering the test to a sample of students.

Kindergarten Readiness by Race/Ethnicity, Gender, and Student Groups

The percentage of students demonstrating readiness for kindergarten increased for most student groups from SY 2022-2023 to SY 2023-2024.

State
44%

Left chart — Race/Ethnicity (2024 / 2023):
- Two or More Races: 52% / 49%
- White: 59% / 56%
- Native Hawaiian or Other Pacific Islander: 39% / 46%
- Hispanic/Latino of Any Race: 24% / 22%
- Black/African American: 40% / 37%
- Asian: 53% / 54%
- American Indian or Alaska Native: 41% / 34%

Right chart — Gender and Student Groups (2024 / 2023):
- Female: 50% / 47%
- Male: 38% / 36%
- Econ. Dis.: 32% / 29%
- SWD: 18% / 17%
- EL: 12% / 10%

Study Overview (1 of 3)

MSDE initiated a study and report to fulfill requirements of the Blueprint implementation.

- The report was to evaluate the current version of the KRA for any racial, cultural, or linguistic biases to ensure the administration of a racially and culturally unbiased statewide kindergarten assessment.

- The results from the administration of the KRA in fall 2022 was evaluated by WestEd and reported on by John's Hopkins's Center for Technology in Education.

- The study looked across groups of Maryland kindergarten students including gender, race/ethnicity, English language proficiency, socioeconomic status, and special education status.

Study Overview (2 of 3)

The study provided findings from the evaluation of racial, cultural, or linguistic bias found in the current KRA.

Differential Item Functioning (DIF) analysis was performed on the KRA items. Demographic information from students is compared and used to determine if groups of students perform differently from each other.

- No DIF found between economically disadvantaged students and non-economically disadvantaged students.
- Every other comparison group (female/male, Black/White, Asian/White, Hispanic/White, EL/non-EL, and IEP/non-IEP) resulted in multiple items with moderate to significant DIF.

- Follow-up evaluation of item difficulties for Hispanic/Latino students and students who are English learners indicate there could be a significant disadvantage to students who are bilingual or multilingual learners.

- The review of the cultural responsiveness of the KRA indicate that some aspects of the current KRA, and in particular some items, need to be reviewed and evaluated for implicit bias toward students from various racial, ethnic, and linguistic backgrounds.

Study Overview (3 of 3)

The report provided recommendations for potential next steps.

- Develop an additional section of the KRA that evaluates Spanish language and literacy proficiency for students whose home or preferred language is Spanish. This section would be in addition to the domains currently in the KRA and should be piloted and field tested prior to administration.

- Convene a Maryland advisory panel of experts and educators from diverse constituencies to review the KRA to ensure that the KRA is responsive to the numerous cultures and backgrounds of Maryland's kindergarten students (blueprint, items, reports, and administration processes and policies).

- Conduct item-review meetings with subject-matter experts and educators to specifically review the items flagged in the study analysis. Revise items, if necessary.

- Review and revise the score reports to ensure that they are interpreted and used appropriately by all intended users, especially teachers and families.

Maryland Kindergarten Readiness Assessment (KRA)

Evaluation of Racial, Cultural, or Linguistic Bias

R 4 K ready for **KINDERGARTEN** MARYLAND

Contents

Executive Summary

Under the Blueprint for Maryland's Future and Education Code § 7-210, the Maryland State Department of Education is tasked with administering a kindergarten readiness assessment to all incoming kindergarten students that is racially and culturally unbiased. This report fulfills requirements in the Blueprint Comprehensive Plan. The Kindergarten Readiness Assessment (KRA) serves to evaluate children's readiness for kindergarten across the essential domains of school readiness. It offers data to local-, regional-, and state-level users to assess children's preparedness for the kindergarten curriculum. This report discusses the KRA's design, development, administration, scoring, reporting, and evidence supporting the reliability and validity of the KRA. Additionally, this report examines potential biases related to race, culture, or language in the KRA, both quantitatively and qualitatively, considering factors like gender, ethnicity, language proficiency, and economic status. Findings related to potential areas of bias and recommended next steps are summarized in the last section of this report.

KRA results provide a measure of children's abilities and skills aligned with the expectations set by the Maryland early learning standards for the end of prekindergarten. The KRA assesses children's readiness via four domains (Language and Literacy, Mathematics, Physical Well-Being and Motor Development, and Social Foundations) via selected-response, performance-task, and observational-rubric item types. The individual student report (ISR) for families, which is generated for every child who is administered the KRA, summarizes each student's results, overall and within each domain, and is available in multiple languages, including Spanish, French, and Chinese. The KRA results provide information about children's readiness for the kindergarten-level curriculum, and should be used with other data and information, including feedback from a child's teacher, to make instructional and programmatic decisions. Performance on the KRA should not prevent or prohibit a child from entering kindergarten.

The KRA is based on Maryland's early learning standards and was developed via an iterative, multiyear process that included cognitive interviews, a pilot, and multiple field tests. All items were reviewed at each stage of the development process, which included review meetings with Maryland State Department of Education (MSDE) staff, Maryland educators, and early learning development experts. These reviews determined that the combination of item types and the content are appropriate for children who are in kindergarten. The KRA utilizes the Rasch Model, a widely employed psychometric scoring model in educational assessment programs. It provides both an overall score and four domain-specific scores. The overall score is utilized for categorizing students into three readiness levels: "Emerging Readiness" (indicating the lowest

performance level on the KRA), "Approaching Readiness," and "Demonstrating Readiness" (indicating the highest performance level on the KRA). A complete description of the design and development of the KRA is provided in Section 2 of this report.

Administration of the KRA includes professional development and support for teachers, as well as available supports for diverse learners, ensuring equitable assessment opportunities for students with disabilities and English learners. A description of the technical development of the KReady online system, which supports the administration, scoring, and reporting of the KRA and provides an interactive platform for educators to manage assessment-related tasks, is also included in Section 2 of this report.

Evidence of validity and reliability for the KRA, based on the Standards for Educational and Psychological Testing (AERA et al., 2014), is described in detail within Section 3 of this report.

Validity refers to the degree to which evidence and theory support the interpretation of assessment scores for proposed uses of assessments. Various aspects of the KRA, including its design, content specifications, item development, and administration procedures, provide evidence of validity. The evidence supporting the KRA content is obtained through the alignment of the KRA items with Maryland's early learning standards and the expert evaluation of item specifications. Cognitive interviews, a pilot, and field tests further contributed to the KRA's validity. Extensive reviews were conducted to ensure the appropriateness of items for diverse student populations. These development steps and review processes aimed to confirm that students engaged with and responded to items as intended. Internal structure evidence is based on classical item statistics, including item difficulty, score-point distribution, and item-total correlations. These statistics were found to be within acceptable ranges and consistent across KRA administrations for all student groups. There are also strong correlations between overall KRA scores and domain scores, supporting the KRA design and inclusion of the essential domains. Additional studies explored the relationship between KRA scores and Maryland state assessment scores in grade 3, showing a moderate, positive association between the two assessments. Further, students who achieved Demonstrating Readiness (the highest level) on the KRA were more likely to meet or exceed expectations on the grade 3 assessments in mathematics and reading.

Reliability measures the consistency of students' scores. High values of Cronbach's alpha indicate that items are closely related, and that students score consistently across items. Further, early childhood educators administering the KRA must undergo training, including the completion of a simulator and content assessment, to support item score reliability.

Differential item functioning (DIF) analysis and its application to the KRA items is also summarized in this report. DIF refers to a statistical process that highlights differences in the probability of individuals from various student groups (e.g., gender, ethnicity, language proficiency) to answer a specific assessment item correctly, conditioned on similar underlying abilities. DIF results can highlight potential concerns about fairness and validity, as it provides

an indication of items that might be biased toward specific student groups. However, it is important to emphasize that DIF results must be substantiated with a qualitative review by subject matter experts to determine if items are genuinely biased.

DIF analyses were conducted on the KRA items for various demographics, including gender, race/ethnicity, socioeconomic status, English learner status, and special education status. The KRA items showed negligible DIF for students who are economically disadvantaged. For gender comparisons, several items displayed moderate DIF in favor of female students. Racial/ethnic DIF comparisons revealed various items with DIF. Items in the Language and Literacy (LL) domain displayed DIF for items related to higher-level language skills. For students with IEPs, many Physical Well-Being and Motor Development and Social Foundations items displayed moderate to significant DIF. DIF in the Physical Well-Being and Motor Development and Social Foundations domains have potential explanations related to range restriction in scores. Detailed results of the DIF analysis are detailed in Section 4 and in the Appendix.

Cultural responsiveness recognizes the profound influence of culture on social interactions and education, encompassing explicit and implicit cultural elements and acknowledging how personal experiences shape knowledge and behavior and influence instructional methods and student development, including assessments. Culturally responsive assessment allows students to demonstrate their mastery from their cultural perspective, promoting equity and fairness. Utilizing the windows-and-mirrors metaphor, culturally responsive assessments enhance engagement and promote equity, allowing students to incorporate their cultural backgrounds into the demonstration of their knowledge, skills, and behaviors. Five design principles for culturally responsive assessments and a set of questions for evaluation of the cultural responsiveness of the KRA are discussed in detail in Section 5.

Conclusions and next steps are outlined in Section 6 of this report. Overall, the results from the DIF analysis and the review of the cultural responsiveness of the KRA indicate that some aspects of the current KRA, and particularly some items, need to be reviewed and evaluated by subject-matter experts for any implicit bias toward students from various racial, ethnic, and linguistic backgrounds. These items should be evaluated by subject-matter experts in additional item-review meetings, where educators can decide whether the items need to be revised to best serve the diverse needs and cultural backgrounds of Maryland's kindergarten students.

Further, the DIF analysis and the follow-up evaluation of item difficulties for Hispanic/Latino students and students who are English learners showed that there could be a significant disadvantage to students who are bilingual or multilingual learners, specifically in their ability to access some items from the Language and Literacy and Mathematics domains. There is strong evidence to support that assessing bilingual or multilingual students in their home language and in English provides the most accurate estimates of their overall ability level (Pitoniak et al., 2009; Durán et al., 2022). Therefore, we also recommend the development of a KRA section that assesses bilingual or multilingual children's skills in their home or preferred language to

ensure that more actionable information can be provided to educators and families to better support these students' learning needs.

Given that roughly 23% of Maryland's public-school students speak Spanish at home, we strongly recommend that MSDE develop KRA sections to be administered in Spanish, in addition to the current version in English. While this will require that MSDE establish additional KRA policies to support decisions about which children to administer the Spanish-language version to, the development of a KRA section to assess bilingual or multilingual children in their home or preferred language would greatly benefit Maryland's kindergarten teachers and students by providing more accurate insight into their abilities.

One approach to ensuring the cultural responsiveness of the KRA is to emphasize the "shared power" principle (Walker et al., 2023). This could be accomplished by convening an advisory panel that includes representatives from MSDE; representatives from various school districts, jurisdictions, and governmental agencies across Maryland; and community members from diverse constituencies, specifically those that represent students from historically marginalized groups. This advisory panel could then be engaged in all phases of the KRA's review and revision process, starting with an evaluation of the current KRA blueprint, items, and reports. Such an intentionally forged partnership that focuses on the experiences of diverse interest groups and ensures all voices are heard throughout the review and revision process is key to achieving cultural validity of the KRA.

In response to the review and findings detailed in this report, the following next steps are recommended:

- Convene an advisory panel of experts and educators from diverse constituencies across Maryland (e.g., MSDE staff, Maryland educators, staff from state universities, representatives of state/local government agencies, and community members) to review the KRA blueprint, items, reports, and administration processes and policies to ensure that the KRA is responsive to the numerous cultures and backgrounds of Maryland's kindergarten students.

- Conduct item-review meetings with subject-matter experts and educators to specifically review the items flagged in the DIF analysis and to determine whether these items require revisions to ensure that they are free from any potential bias. Revise items, if necessary.

- Develop an additional section of the KRA that evaluates Spanish language and literacy proficiency for students whose home or preferred language is Spanish. This new Spanish-language section should be piloted and field tested prior to implementation within the operational KRA. Further, this additional section should be administered in

addition to the domains currently included in the KRA. This would also require that MSDE establish policies to support the administration of this additional section.

- Review and revise the score reports to ensure that they are interpreted and used appropriately by all intended users and constituents, especially teachers and families.

1 Introduction and Purpose of This Report

The purpose of the Kindergarten Readiness Assessment (KRA) is to provide valid and reliable information on children's learning and development across the essential domains of school readiness.[1] This information can be used by relevant users at the local, regional, and state levels to better understand children's preparedness for kindergarten. Detailed score reports at the individual, classroom, school, district, and state levels inform policy, research, and programmatic decisions, and families can learn about each child's skills, knowledge, and developmental needs.

In this report, we first provide an overview of the design and development of the KRA, current accommodations and supports for students with disabilities and students who are English learners, current scoring and reporting procedures, evidence of reliability and validity of the KRA, and an evaluation of the cultural responsiveness of the KRA. A broader purpose of this report is to evaluate the current version of the KRA for any racial, cultural, or linguistic biases. In addition, the results from the administration of the KRA in fall 2022 were quantitatively evaluated for differential item functioning (DIF) across groups of Maryland kindergarten students (as defined by gender, race/ethnicity, English language proficiency, socioeconomic status, and special education status). The results from the DIF analysis and the review of the cultural responsiveness of the KRA, including any implications of racial, cultural, or linguistic biases in the current KRA, are summarized in this report. Lastly, this report provides recommendations for next steps that could be taken to ensure that the KRA is culturally responsive and to minimize any potential biases in the KRA.

---

[1] The U.S. Department of Education defines the essential domains of school readiness as language and literacy development, cognition and general knowledge (including early mathematics and early scientific development), approaches toward learning, physical well-being and motor development, and social and emotional development.

2 KRA Design and Development

The KRA was developed by the Maryland State Department of Education and the Ohio Department of Education, in conjunction with the Johns Hopkins University Center for Technology in Education (JHU CTE), WestEd, and a Technical Advisory Committee (TAC) composed of early learning and assessment experts.

The first versions of the KRA, referred to as KRA 1.0 and KRA 1.5, were developed with Race to the Top – Early Learning Challenge grants that were awarded to Maryland and Ohio, beginning in 2011. The KRA 1.5 was the operational version of the KRA in Maryland from 2014 to 2017. A revised and improved version of the KRA, referred to as KRA 2.0, was developed with an Enhanced Assessment Grant awarded to Maryland by the U.S. Department of Education, beginning in 2013. Several states partnered with Maryland during the development of the KRA 2.0, including Connecticut, Indiana, Michigan, Ohio, and Tennessee. The KRA 2.0 was first administered operationally in fall 2018.

Unless specifically stated otherwise in this report, any references to the KRA represent KRA 2.0, as it is the current version administered in Maryland.

2.1 KRA Content Standards

The KRA is a criterion-referenced assessment based on the Maryland early learning standards, which include the domains, strands, standards, and essential skills and knowledge that form the basis of the KRA. The KRA is based on prekindergarten standards and incorporates the essential domains of school readiness as defined by the U.S. Department of Education.[2] A detailed description of the early learning standards and development process is presented in the *KRA 2.0 Development and Technical Report* (WestEd, 2018).

2.2 KRA Item Types

A KRA item is one question or observation that aligns to a specific early learning standard and that results in one recorded score. The KRA includes three item types: selected response, performance task, and observational rubric.

---

[2] The Social Foundations domain for the KRA incorporates the essential domains of social and emotional development and approaches toward learning.

Selected-response items consist of a question or prompt and three possible answer options, of which there is only one correct answer. These item types include materials that have student-facing colorful images and paper-based manipulatives. A student indicates his or her response by touching one of the three answer options. Selected-response items are worth one score point. The benefits of selected-response items are that they require the least amount of time to administer and can be administered via the KRA App.

Performance-task items consist of an activity or action that the student completes in response to a prompt. In some instances, manipulatives are provided with performance tasks, which allow the student to demonstrate the skill being assessed. Performance-task items are scored with a rubric that is based on the proficiency of the student's performance, and are worth one, two, or three score points. The benefit of performance-task items is that they allow a student to demonstrate his or her knowledge and, in some instances, to provide an explanation or reason. Some performance-task items can be administered via the KRA App.

Observational-rubric items describe specific behaviors or skills that a student should demonstrate during typical classroom activities. The teacher evaluates and scores each student's behaviors or skills, using a rubric that describes the quality for each criterion. Observational-rubric items do not require the teacher and the student to directly interact (i.e., the student is unaware of the teacher's intention to assess) and, therefore, provide the advantage of assessing the student in their natural environment (both inside and outside the classroom, such as playground, cafeteria, etc.).

2.3 KRA Blueprint

The KRA Blueprint, shown in Table 2.3.A, outlines the distribution of selected-response (SR) items, performance-task (PT) items, observational-rubric (OR) items, total items, total points, and percentage of total points across the domains.

Table 2.3.A. KRA Blueprint

Domain	SR	PT	OR	Total Items	Total Points	Percentage of Total Points
Language and Literacy	7	6	4	17	33	35%
Mathematics	2	11	0	13	22	23%
Physical Well-Being and Motor Development	0	0	9	9	18	19%
Social Foundations	0	0	11	11	22	23%
Total	**9**	**17**	**24**	**50**	**95**	**100%**

2.4 KRA Development Process

The development of the current KRA included standards alignment, cognitive interviews, a pilot, and two field tests. Each step of the process iteratively informed subsequent activities, providing critical evidence to support the validity and reliability of the KRA for its intended purpose. A summary of the development process is provided in the subsequent subsections of this report; however, all details about the KRA development process can be found in the *KRA 2.0 Development and Technical Report* (WestEd, 2018).

2.4.1 KRA Cognitive Interviews

WestEd, in conjunction with the Connecticut Office of Early Childhood, conducted cognitive interviews to inform the initial development of the KRA. The cognitive interviews were conducted in November 2014 with kindergarten students and their teachers from 10 school districts in Connecticut. As a result of the standards alignment, WestEd identified potential new or modified early learning standards in the domains of Language and Literacy and Social Foundations. Maryland educators and students did not participate in the cognitive interviews, as Maryland was actively administering the operational KRA 1.5.

This cognitive interview study was designed to evaluate items that were developed to assess the new or revised standards and to provide information on the students' engagement with the newer items. A total of 22 items, including selected-response, performance-task, and observational items, were evaluated for inclusion in the KRA. Items from the previous version were also administered via the KRA App to allow WestEd researchers to observe students' interaction with the items on a tablet.

The cognitive interviews were conducted with 55 kindergarten students and 10 teachers from a range of public-school districts (urban, suburban, and rural). Teachers and students participated voluntarily. At each school, teachers assisted in the selection of students to participate, based on a representative sample from each class. Students with limited understanding of English or severe cognitive disabilities were not selected to participate in the cognitive interviews. Given the age of the students in this study, a retrospective inquiry process (Sato et al., 2010) was used to be developmentally appropriate. Rather than pausing and asking probing questions between each item, the interviewer encouraged retrospection when students were naturally inclined to express their thoughts as they went through the assessment. All interviews were conducted in the students' schools during school hours and lasted approximately 10–20 minutes. After the student interviews were completed, teachers were asked to share their general observations about the student interviews and to provide feedback on each item observed during the student interview process. Additionally, teachers were asked to read and provide feedback on the observational items.

Findings from the cognitive interviews indicated that students were generally able to access all of the items and responded as required during the cognitive interviews. In general, the students did not appear anxious or overwhelmed by the items, and they reacted positively to the one item that utilized manipulatives. Students also expressed working knowledge of a tablet, and they reacted positively to the KRA App items. Some items, where the context presented in the item was contrary to the current situation (e.g., when an item graphic about weather did not match the weather outside), caused confusion for several students. In a few items, the directions were not explicit enough to elicit the desired response (e.g., "Touch the place where I should start reading"); however, rephrasing the prompt to be more precise (e.g., "Touch the word . . .") often resolved the student's confusion. When asked about the difficulty of items, the students tended to rate items as "easy," regardless of how they performed on the items. The distribution of score points and the item difficulty ratings for the selected-response and performance-task items indicate that most items performed as expected.

The teachers' impressions of their students' performance provided a lens through which the students' responses and comments could be appropriately interpreted. Teachers attributed the fact that many students tended to rate the items as easy to a general lack of performance anxiety, rather than to the actual difficulty of the items. Most teachers agreed that, overall, students performed as expected. However, some teachers reported being surprised by the performance of specific students. For example, some students surprised their teachers by being more comfortable and confident in their responsiveness than the teachers thought they would be.

For most of the paper-based items, teachers agreed that the graphics and directions seemed appropriate for and accessible to students entering kindergarten. However, they reported that they would not expect students entering kindergarten to appropriately use the grammar rules associated with subject-verb agreement and verb tenses, and that items related to these skills were not likely to provide actionable data that would inform their instruction. Teachers did not provide much specific feedback on accessibility for English learners or students with disabilities. They expressed the belief that items would be more difficult for English learners to understand because of issues with receptive language comprehension, or that students with disabilities might struggle with some items, depending on the specific nature of their disability. They did not report any observable bias of the items.

Multiple teachers remarked favorably about the inclusion of items aligned to standards in the Social Foundations domain. They reported that social and emotional skill building is a major focus of instruction in their kindergarten classrooms. They expressed some excitement that the inclusion of these items might help educators place more appropriate value on the time and instruction that teaching these skills really requires. For most of the observational items, teachers agreed that they would be able to observe students easily for the purposes of assigning a rating on the rubrics (i.e., they would have multiple opportunities and that it would not disrupt their normal routine). An issue that came up in several interviews was that teachers

do not always see an inherent value in using social-emotional rubrics. They explained that, as experienced teachers, they intuitively get a sense of their students' social and emotional development, and that rating students on the observational rubric would not really provide additional information for classroom instruction.

Teachers were generally excited to see students using the KRA App items and reported that they were pleased that the assessment utilized mixed media (some paper items, some manipulatives, and some tablet items). They also reported that they liked the idea of the tablet providing them with some relief in terms of materials management. Teachers were also impressed by students' level of engagement with the KRA App items. On the other hand, they noticed that impulsiveness was a critical factor in how students performed on these items. They suggested that experienced tablet users, which described most students, were likely to be inclined to try to complete the "games" as fast as they could, and they expressed some concern that the "gamer" mentality prioritized speed over finding the correct answer.

2.4.2 KRA Pilot

The feedback from teachers and student responses to items during the cognitive interviews were used to make revisions to the KRA. These include specific feedback about the alignment of items to the standards, the need to pay close attention to the situational context of items in relation to potential "real-world" situations during times of administration, the clarity and specificity of directions, and the potential for paper-based items to be adapted as KRA App items. Then, a pilot was conducted to evaluate the utility of the revised KRA items. The pilot sample consisted of 62 teachers from 37 schools in 18 school districts. Each participating teacher was asked to choose five students that represented the diversity of students in their classroom, which consisted of 264 total students. Teachers were also advised to provide accommodations for students with disabilities that were consistent with current state assessment policies and guidelines. Although the pilot was conducted only in Connecticut, the results informed improvements to the quality and efficiency of the KRA for all participating states. The pilot comprised three major components: teacher training, administration and review of assessment items, and a teacher survey. Table 2.4.2.A summarizes each component.

Table 2.4.2.A. KRA Pilot Components

Component	Description	Estimated Time
Teacher Training	Teachers attended a training session via webinar or in person. Attendance at one session was required, but teachers had the option to attend both.	90 minutes
Administration and Review of Assessment Items	Upon completion of the training, teachers administered an assigned section of selected-response and performance-task items to five students and entered the students' responses to items into the online system; teachers also reviewed the observational-assessment items but did not administer these items—they provided feedback on the items in the survey.	15–20 minutes per student for selected-response and performance-task items 15–20 minutes for review of observational items
Teacher Survey	Teachers filled out an online survey at the end of the pilot administration window. The teacher survey was designed to collect general information and item-specific feedback from the teachers.	10–15 minutes

Note. **The estimated administration times were provided to assist teachers in planning; however, the pilot was not timed.**

Most of the teachers strongly agreed or agreed that the training session sufficiently prepared them for the pilot (37% strongly agreed and 49% agreed), that the technology system was easy to navigate (11% strongly agreed and 65% agreed), and that they learned valuable information about the students in their classrooms during the pilot (19% strongly agreed and 61% agreed).

For the selected-response and performance-task items (within the Mathematics and the Language and Literacy domains), the majority of teachers strongly agreed or agreed that the items can be easily integrated into their regular classroom activities, are accessible for English learners, and are accessible for all students (with allowances or accommodations); that the scoring information was clear and easy to apply; and that the images and/or manipulatives in the items were clear and easy for students to understand. In addition, most of the teachers believed that these items were best administered individually, rather than in small groups or to the whole class.

For the observational items (within the Language and Literacy, Social Foundations, and Physical Well-Being and Motor Development domains), the majority of the teachers strongly agreed or agreed that the skills and behaviors in the observational items would be observable during the first six weeks of school, that the skill or behavior would be easy to observe in situations that occur within their normal school day, that all of the students could be rated on the skill or behavior (with appropriate allowances or accommodations), and that the three level descriptors within the scoring rubrics (i.e., Evident, In Progress, and Not Yet Evident) were easily distinguishable.

Overall, the pilot data and findings supported the assumptions underlying the assessment design. The feedback received from the teachers via the survey endorsed the design of the KRA. The results of the pilot further informed the KRA item development process and allowed the states to review and finalize the KRA Blueprint (as shown in Section 2.3). Item enhancements focused on maximizing accessibility for English learners and students with disabilities.

2.4.3 KRA Field Tests

A field test of the KRA items was conducted in Maryland, Ohio, and Tennessee during fall 2016.[3] A set of 85 items were field tested (55 items were selected-response or performance-task items and 30 items were observational-rubric items). Because the item types and administration procedures were familiar to teachers in Maryland and Ohio, their field test–specific training consisted of a one-hour webinar. In addition, they received guidance from representatives of their state departments of education via email and had access to other support materials (e.g., a recording of the webinar) via the online system. JHU CTE provided a more substantive training to educators in Tennessee via a two-day, in-person, training-of-trainers session. The trainers who attended the JHU CTE training-of-trainers session were responsible for training and supporting the teachers in Tennessee. JHU CTE also held a webinar for data managers in Tennessee, so that the online system could be prepared for data collection.

A total of 243 teachers (53 in Maryland, 77 in Ohio, and 113 in Tennessee) and 3,876 kindergarten students participated in the field test. Teachers were recruited to participate in the field test by state department staff. To ensure a balanced representation of students, state staff members attempted to recruit a representative convenience sample from across their states' general kindergarten populations, with purposeful sampling to include low-socioeconomic-status schools, English learners, and students with IEPs.

The field test forms were distributed at the school level, so that the same form was administered by every teacher within a participating school. In Maryland, teachers

---

[3] The initial planned outcome of the field test in 2016 was to use the findings to select the final set of items for one operational KRA 2.0 form, to be administered in fall 2017. However, in December 2016, the states decided to amend the scope and timeline of the Enhanced Assessment Grant project so that multiple KRA 2.0 forms, equated to each other and to KRA 1.5, could be created for use beginning in fall 2018. This change in scope allowed for administration of the KRA 2.0 without any delays in the reporting of results in 2018 and beyond, and therefore led to an additional field test in fall 2017.

administered the field test items after the completion of the operational KRA 1.5 administration. In Ohio, teachers administered the field test throughout the operational KRA 1.5 administration window. In Tennessee, teachers administered the field test items only, and did not administer the operational KRA 1.5. Participating teachers were asked to provide feedback about their experience via an online survey. Responses were collected throughout the administration window in each state. The survey consisted of questions related to the field test materials, the field test items, and the overall experience. It also included state-specific questions that were created by state departments for their own analyses.

Because the scope of the Enhanced Assessment Grant project was amended (see prior footnote), the KRA 2.0 items could be field tested again in fall 2017 in Maryland and Ohio. This additional year of field testing, analogous to an embedded field test model, allowed two forms of the KRA 2.0 items to be psychometrically scaled with the KRA 1.5 items, and equated with each other. In fall 2017, Maryland and Ohio field tested the 78 KRA 2.0 items, in conjunction with their operational administrations of the KRA 1.5; the KRA 2.0 items were distributed across 16 forms and "appended" to the existing KRA (i.e., KRA 1.5), in lieu of the typical embedded field test model. All students in Maryland and Ohio who were assessed with the KRA 1.5 in fall 2017 participated in the field test. The 16 field test forms were distributed at the school level, so that the same form was administered by every teacher within a participating school.

Classical item analyses, item calibrations, scaling, and form equating analyses were conducted after both field test administrations to ensure that the items functioned as expected, resulting in the development of two new KRA forms (KRA 2.0 Forms A and B). Currently, KRA 2.0 Form A is administered throughout Maryland (and South Carolina and Hawaii); KRA 2.0 Form B has not been utilized within Maryland (or any other state).

2.5 KRA Scoring and Reporting

Given that the KRA includes a sample of items that can be used to measure readiness for kindergarten, percent-correct scores would not provide a complete explanation of a student's readiness for kindergarten. Instead, raw scores (i.e., the total score points obtained across all items) on the KRA are converted to scale scores. Scale scores account for the difficulty of individual items and forms, providing consistency in the interpretation of results and allowing for comparison of results across cohorts and forms.

2.5.1 KRA Scoring

The KRA utilizes the Rasch model to define the relationship between the assumed latent trait (readiness for kindergarten) and the probability of a student correctly answering a given KRA item. This model assumes that responses are a function of a student's knowledge about the assessment content and of the difficulty of the item. This model allows the student's score and the difficulty of the item to be placed on the same scale, known as theta (θ), which represents

the latent trait being measured. This θ scale allows for direct interpretation of the difficulty of an item and the probability of a student answering an item correctly. The probability that a student will answer a question at a given level is determined by whether the student's score is below, at, or above the difficulty threshold for that level.

The KRA items were calibrated using WINSTEPS measurement software. A more detailed description of the KRA scaling process, including the item parameters and fit statistics, can be found in the *KRA Technical Report* (WestEd, 2014) and the *KRA 2.0 Development and Technical Report* (WestEd, 2018).

The θ scale is centered at 0 and extends in both positive and negative directions. Applying a linear transformation to the θ scale is desirable because it allows for a scale that is more easily understood by all intended users and that does not include negative values. The θ scores determined by Rasch scaling are converted using a linear transformation such that $scale\ score = 12 * \theta + 250$. The KRA scale is truncated at θ scores of ±4, which results in minimum and maximum scale scores of 202 and 298, respectively.

2.5.2 KRA Reporting

The KRA overall scale score determines each student's performance level: Demonstrating Readiness, Approaching Readiness, or Emerging Readiness. Table 2.5.A shows the performance levels and their descriptions, including their associated overall score ranges.

Table 2.5.A. Performance Levels and Overall Scale Score Ranges for the KRA

Performance Level	Description	Overall Scale Score Range
Demonstrating Readiness	The child demonstrates foundational skills and behaviors that prepare him/her for curriculum based on kindergarten standards.	270–298
Approaching Readiness	The child demonstrates some foundational skills and behaviors that prepare him/her for curriculum based on kindergarten standards.	258–269
Emerging Readiness	The child demonstrates minimal foundational skills and behaviors that prepare him/her for curriculum based on kindergarten standards.	202–257

To show relative strengths in each student's performance, domain scale scores, based on the subset of KRA items that are aligned to each domain, are also reported for each student. The domain scale scores are reported using the same scale as the overall score (except for Physical Well-Being and Motor Development, which has a lower maximum score due to the limited number of items and score points for that domain). Table 2.5.B shows the ranges of possible scale scores for each domain. Caution must be taken when interpreting domain scale scores, as these scores are determined by a subset of the items that compose the entire KRA, meaning that they provide a less-precise measure of ability.

Table 2.5.B. Domain Scale Score Ranges for the KRA

Domain	Scale Score Range
Language and Literacy	202–298
Mathematics	202–298
Physical Well-Being and Motor Development	202–293
Social Foundations	202–298

Note. **The Physical Well-Being and Motor Development domain has a lower maximum score due to limited numbers of items and score points within the domain.**

Upon completion of the KRA, each student receives an individual student report (ISR), which can be generated by the teacher upon completion of the assessment with the student.[4] The ISR provides the student's overall score and associated conditional standard error of measurement; performance level, based on the overall score; domain scores and associated conditional standard errors of measurement; and completion status. The ISR is available in English, Spanish, French, and Chinese.

In addition to the ISR, multiple reports are available to teachers via the KReady system. The following reports can be generated by teachers throughout and after the KRA administration window:

- *Interactive Data Displays*: The Interactive Data Displays are interactive charts and graphs that present the KRA data in multiple ways, including the option to filter by student group.

---

[4] A sample KRA ISR is available at: https://www.marylandpublicschools.org/about/Documents/DAAIT/KRA/KRAISR.pdf

- *Domain Data Export*: This report is a Microsoft Excel file of a teacher's class roster, organized by domain, showing total raw points earned by each student.

- *Data Results Export*: This report is similar to the Domain Data Export but is organized by item. The spreadsheet can be sorted and filtered to meet the teacher's needs.

- *Class Item Results*: This report is a PDF with scoring rubrics, showing student performance by item.

- *Individual Student Item Results*: This report is a PDF of student scores by item, including scoring rubrics. This report can be printed separately for each student, showing the student's scores for all items or only for selected items.

The KReady system also offers a variety of reports for school and district administrators. Having access to the KRA data and results allows school and district administrators to provide targeted instructional support to programs and individual children. In addition to the previously described Interactive Data Displays and Domain Data Export reports, the following reports can be generated by school and/or district administrators:

- *KRA ISR Report*: This report is a Microsoft Excel file that includes all student data (including demographic information), students' overall and domain scores, and students' item-level scores. This report also includes links to view students' ISRs.

- *KRA Percentage Completion Report*: This report provides the percentage of students in a school or district who have completed the KRA.

- *KRA Completion by Item Report*: This report provides the KRA items that have been completed for each student in a school or district.

- *ISR ZIP File*: This ZIP file includes all student ISRs.

2.5.3 Interpretations of KRA Scores

The KRA is designed to provide each student with an overall scale score, a performance level (Emerging, Approaching, or Demonstrating Readiness) based on the overall scale score, and domain scale scores. The overall scale score and performance level support annual state-level reporting, as reported in the annual Readiness Matters reports.[5] The domain scale scores are included on the ISR to describe the student's relative strengths and areas of need. The ISR also includes ideas for ways in which family members can support their child's progress and converse with their child's teacher about their child's progress. The additional reports described in Section 2.5.2, including the interactive data displays in the KReady system, provide detailed information that allows teachers to explore results for individual students and groups of students.

---

[5] These reports are available at: https://earlychildhood.marylandpublicschools.org/kindergarten-readiness-report

Scale scores should be used to support the intended purposes of the KRA, which are to identify children's readiness for the kindergarten curriculum and to identify students' strengths and areas of need. KRA results should not be used for any other purposes. MSDE and JHU CTE provide resources to teachers and administrators to support the appropriate use of KRA scores and reports, with the intention to mitigate the use of KRA scores for unintended consequences.

2.6 KRA Professional Development and Administration

Led by JHU CTE, the professional development to support the KRA administration is one aspect of overall implementation that requires intentional design, customized development, and delivery of information about the assessment and technology systems. In addition, effective implementation includes careful attention to the needs of the state, as well as to the technical requirements that result in personalized approaches to professional development and support for all who interact directly with the KRA.

As part of its development of training content and implementation approach, JHU CTE

- learns about Maryland's unique needs, policies, and processes to scale implementation effectively;
- offers professional development through a variety of formats, including online communities, to engage relevant audiences and to promote resource sharing;
- collaborates with MSDE staff to ensure that the professional development effectively supports the system;
- enhances professional development content regularly to communicate updates to the assessment and policy information; and
- implements a multilevel evaluation strategy, including simulation technology, surveys, and fidelity checklists, to promote training and assessment implementation fidelity.

For each administration of the KRA, the professional development is refined, enhanced, and expanded in collaboration with state leadership based on ongoing evaluation and feedback from teachers, trainers, and local leaders. The professional development acknowledges that users will have varying levels of knowledge and experience with the KRA and provides differentiated supports to accommodate these variations.

2.6.1 KRA Professional Development

In addition to teacher professional development, training is provided to district data managers, district- and school-level administrators, and other local leaders, unique to the needs of Maryland. These training sessions are interactive, hands-on, and designed to provide participants with opportunities for practice and reflection.

The comprehensive professional development approach includes

- face-to-face and webinar sessions for district and other local leaders, as well as regional trainers, to prepare for the rollout of the KRA each year and to review the technology enhancements and updates;

- webinar training for district data managers, coupled with video resources, to provide a detailed walkthrough of the KReady system and data manager functionality;

- face-to-face and webinar touchpoint sessions with trainers to ensure that they remain abreast of technology, updates, and areas of support needed in the field; and

- webinar trainings for school and district administrators for understanding the KRA, supporting teachers, and interpreting and using KRA data.

The ongoing feedback loop through research and evaluation reflects JHU CTE's commitment to the continual improvement and enhancement of the professional development related to the assessment and technology system to meet the varied and changing needs of the various KRA users.

2.6.2 KRA Supports for Diverse Learners

All students, including students with disabilities and students who are English learners, are required to participate in the KRA, and their results are a part of the summary reports. A fully accessible approach to assessment design and implementation is necessary to ensure that students with diverse learning characteristics have the opportunity to demonstrate their knowledge and skills. At the same time, MSDE needs to be confident in the reliability of results obtained from the KRA when conducting analyses and making policy decisions. In addition, school administrators, teachers, and families need to be knowledgeable about where their students are functioning developmentally to identify focus areas for instruction that promote growth in individual students. The guidance document for administering the KRA to diverse populations of students is referred to as the *Guidelines on Allowable Supports for the Kindergarten Readiness Assessment*.[6]

In recognition of the need for practitioners to fully understand the importance of differentiation of administration to meet the needs of diverse learners, a specific portion of professional development and training is devoted to instruction on these guidelines. These guidelines provide detailed information on the strategies and practices that support a differentiated administration of the assessment. The KRA training ensures that trainers and teachers will learn about the Universally Designed Allowances that are available for all students, including materials presentations, procedures, and settings that can be used to ensure that all students can access KRA items. These guidelines also provide an item-by-item decision-making process for providing supports to students with disabilities and to English learners, called Level

---

[6] These guidelines are available at: https://pd.kready.org/data/ck/sites/116/files/MD_KRA%20Guideline%20final%20rv.pdf

the Field supports. These supports provide equal access and opportunity for all students to participate in the KRA, without substantially altering what a student is expected to do. They are intended to reduce or even eliminate the effects of a student's disability or limited English proficiency. Through in-depth review, practice, demonstration, and reflection, the professional development ensures that participants understand how to appropriately administer the KRA to diverse populations of students.

2.6.3 Ready for Kindergarten Online System

Technical development of the Ready for Kindergarten (KReady) online system is led by JHU CTE. The KReady online system was developed based on requirements gathered from key personnel in the state departments of education and WestEd.[7]

The KReady online system supports

- administration and scoring of the KRA;
- transfer of data to and from state longitudinal data systems;
- generation of reports summarizing student-level results;
- generation of individual student reports (ISRs) for parents and families;
- generation of reports to monitor completion of the KRA for key personnel at the local, district, and state levels;
- management of teacher and student enrollment information;
- management of the assessment content and supporting materials; and
- implementation of professional development.

In addition, teachers have the option of using the KRA App to administer a subset of the KRA items directly to the student with a tablet, eliminating the need for data entry altogether for those items. Through the KRA App, each item is read aloud to the student, and the student independently works through the item. The score for each item is automatically incorporated into the KReady online system, thereby saving teacher administration time.

---

[7] More details about the KReady system and the professional development supports can be found at: https://pd.kready.org/105956

3 Evidence of Validity and Reliability for the KRA

The *Standards for Educational and Psychological Testing*, published by the American Educational Research Association, the American Psychological Association, and the National Council on Measurement in Education (AERA et al., 2014), provide detailed explanations of validity and reliability. These standards were used to guide the entire design, development, scoring, administration, and reporting processes for the KRA. The statistics presented throughout this section are based on data collected during the KRA administration in fall 2022.

3.1 KRA Validity

According to the *Standards for Educational and Psychological Testing*, "validity refers to the degree to which evidence and theory support the interpretation of assessment scores for proposed uses of assessments." Further, "the process of validation involves accumulating relevant evidence to provide a sound scientific basis for the proposed score interpretations"; therefore, "statements about validity should refer to particular interpretations for specified uses" (AERA et al., 2014, p. 11).

Every aspect of an assessment, including its design, content specifications, item development, psychometric characteristics, and administration procedures, provides evidence in support of its validity (or evidence of lack of validity). Therefore, every section of this report provides evidence of validity for the use of the KRA to describe children's preparedness for kindergarten curricula.

3.1.1 Evidence Based on Test Content

The KRA Blueprint, item specifications, and item development process provide validity evidence based on test content.

As described in Section 2.1 of this report, the KRA is aligned to Maryland's early learning standards and incorporate the essential domains of school readiness, as defined by the U.S. Department of Education. The KRA Blueprint emphasizes all domains of school readiness and utilizes multiple item types to best assess the skills and behaviors within each domain. Prior to

item development, detailed item specifications aligned to the early learning standards were created by WestEd content experts and reviewed by content experts from the state departments of education. The item specifications ensure alignment to the early learning standards and describe the parameters for item development. As described in Section 2.4 of this report, and in further detail in the *KRA 2.0 Development and Technical Report* (WestEd, 2018), cognitive interviews, a pilot, and a field test were conducted. Each step of these processes further contributed to the validity of the KRA and provided opportunities for expert and educator review and feedback, in addition to statistical analyses.

Prior to field testing, every KRA item went through a bias and content review. The bias and content review committees consisted of early childhood educators from the KRA states. Staff from the state departments of education also reviewed and approved each item prior to field testing. To ensure maximum accessibility for English learners, experts from the WIDA Consortium reviewed and provided feedback on every KRA item prior to field testing. The extensive rounds of review and feedback ensure fidelity to the standards and appropriateness for use with children entering kindergarten.

All students, including students with disabilities and students who are English learners, are required to be assessed. A fully accessible approach to assessment design and implementation was necessary to ensure that students with diverse learning characteristics had the opportunity to demonstrate their knowledge and skills. The guidance document for administering the KRA to diverse populations of students is referred to as the *Guidelines on Allowable Supports for the Kindergarten Readiness Assessment*.[8] These guidelines provide detailed information on the strategies and practices that support differentiated administration of the assessment.

Training on the KRA ensures that teachers learn about the Universally Designed Allowances that are available for all students, including materials presentations, procedures, and settings that can be used to ensure that all students can access the items. These guidelines also provide an item-by-item decision-making process for providing supports to students with disabilities and to English learners. These supports, called Level the Field supports, provide equal access and opportunities for all students to participate in the KRA without substantially altering what a student is expected to do. They are intended to reduce or even eliminate the effects of a student's disability or limited English proficiency.

3.1.2 Evidence Based on Response Processes

The cognitive processes engaged in by test takers and the extent to which the processes of observers are consistent with the interpretation of scores can provide evidence supporting the fit between the construct and the nature of the performance or response that test takers engaged in (AERA et al., 2014). The cognitive interviews described in Section 2.4 of this report and in the *KRA 2.0 Development and Technical Report* (WestEd, 2018) were conducted so that

---

[8] These guidelines are available at: https://pd.kready.org/data/ck/sites/116/files/MD_KRA%20Guideline%20final%20rv.pdf.

the assessment developers could better understand new item types and formats and to confirm hypotheses about access to the aligned content. The cognitive interviews allowed the developers to evaluate assumptions about the intent of an item or task, including the reasoning processes that students used to respond to the item.

In addition to the cognitive interviews, the teacher surveys that were conducted during the pilot and the field tests included questions designed to provide evidence that the students were engaging with and responding to items as intended. As described in the *KRA 2.0 Development and Technical Report* (WestEd, 2018), the results from the teacher surveys include strong evidence to confirm that the response processes of students were consistent with the intended designs of the items.

3.1.3 Evidence Based on Internal Structure

The KRA items are evaluated annually for their mean, standard deviation, difficulty (*p*-value), score-point distribution, and discrimination (item-total correlation). The *p*-value statistic is a measure of item difficulty (or item easiness) and falls between 0 and 1. For polytomous items (i.e., items scored on a rubric), the *p*-value statistic is relative to the maximum item score and is calculated by dividing the mean by the maximum possible score for each item. The score-point distributions provide the percentages of students who received each score point on a specific item. The item-total correlation is used to evaluate item discrimination by determining an individual item's relationship to the overall (or total) score, excluding the item of interest. Item-total correlations are values between −1.00 and 1.00, where 0 represents no correlation.

Table 3.1.3.A provides a summary of the classical item statistics for the KRA in fall 2022. These statistics fall within acceptable ranges. The classical item statistics for all 50 KRA items administered in fall 2022 are provided in the *Maryland KRA: Annual Technical Report* (2022–2023) (WestEd, 2023).

Table 3.1.3.A. Summary of Classical Item Statistics for the KRA in Fall 2022

Domain	No. of Items	p-Value			Item-Total Correlation		
		Mean	SD	Range	Mean	SD	Range
All	50	0.73	0.13	0.36–0.95	0.54	0.12	0.28–0.72
Language and Literacy	17	0.66	0.12	0.36–0.86	0.52	0.15	0.28–0.70
Mathematics	13	0.69	0.15	0.37–0.85	0.49	0.11	0.31–0.66
Physical Well-Being and Motor Development	9	0.86	0.06	0.78–0.95	0.55	0.08	0.43–0.69
Social Foundations	11	0.78	0.07	0.68–0.86	0.61	0.07	0.51–0.72

The overall score and the domain scores for the KRA are also strongly correlated, as evidenced by the Pearson correlation coefficients shown in Table 3.1.3.B.

Table 3.1.3.B. Pearson Correlation Coefficients Between the Overall Score and the Domain Scores

Domain	Overall	LL	MA	PD	SF
Overall	1.00				
Language and Literacy (LL)	0.93	1.00			
Mathematics (MA)	0.85	0.82	1.00		
Physical Well-Being and Motor Development (PD)	0.77	0.60	0.52	1.00	
Social Foundations (SF)	0.81	0.62	0.53	0.81	1.00

Note. N = 59,893.

The detailed descriptions of the item calibration process and reporting scale can be found in the *KRA 2.0 Development and Technical Report* (WestEd, 2018), and the descriptive and reliability statistics that are described in Section 3.2 provide additional validity evidence based on internal structure.

3.1.4 Evidence Based on Relations to Other Variables

Several studies have been completed to explore the relationship between KRA scores and scores from Maryland state assessments at grade 3.

The first of these studies was completed by the Regional Education Laboratory (REL) Mid-Atlantic in December 2019, which explored the validity and reliability of a K–3 school-level growth measure (Dragoset et al., 2019). Dragoset et al. (2019) found that the overall KRA score predicted grade 3 achievement reasonably well, relative to other kindergarten assessments. The correlation between the overall KRA scores and grade 3 PARCC (Math and Reading) scores was 0.53, indicating that KRA scores accounted for approximately 28% of the variance in grade 3 PARCC scores (Dragoset et al., 2019). These results indicate that the KRA has a moderate, positive association with grade 3 achievement, which is reasonable considering that the two assessments were administered almost 3.5 years apart from each other.

As a follow-up to the REL Mid-Atlantic study, WestEd further explored the relationship between the KRA scores and grade 3 PARCC scores. Utilizing data from two cohorts of Maryland students who were administered both the KRA and grade 3 PARCC assessments (KRA 2014/PARCC 2018 and KRA 2015/PARCC 2019), WestEd compared the percentage of students at grade 3 PARCC performance levels (Levels 1, 2, 3, 4, and 5) for each KRA performance level (Emerging, Approaching, and Demonstrating Readiness).

Table 3.1.4 A and Table 3.1.4.B summarize the percentage of students at grade 3 PARCC performance levels for each KRA performance level for PARCC Reading and PARCC Math, respectively. As shown in these tables, students who achieved Demonstrating Readiness on the KRA were more likely to achieve PARCC Levels 4 and 5, and students who achieved Emerging Readiness on the KRA were less likely to achieve PARCC Levels 4 and 5.

Table 3.1.4.A. Percentage of Students at Grade 3 PARCC Reading Performance Level for Each KRA Performance Level

	PARCC Level 1	PARCC Level 2	PARCC Level 3	PARCC Level 4	PARCC Level 5	PARCC Level 4/5	Total
KRA Emerging Readiness	50.7	22.7	16.5	9.8	0.3	10.1	14,781
KRA Approaching Readiness	24.0	22.1	24.7	27.7	1.5	29.2	35,824
KRA Demonstrating Readiness	6.9	10.7	20.3	53.3	8.8	62.1	47,025
Total	19,347	16,316	20,815	36,454	4,698	41,152	97,630

Note. **Results are from the KRA 2014/PARCC 2018 and the KRA 2015/PARCC 2019 cohorts. PARCC levels were defined as: 1 = Did Not Yet Meet Expectations, 2 = Partially Met Expectations, 3 = Approached Expectations, 4 = Met Expectations, and 5 = Exceeded Expectations.**

Table 3.1.4.B. Percentage of Students at Grade 3 PARCC Math Performance Level for Each KRA Performance Level

	PARCC Level 1	PARCC Level 2	PARCC Level 3	PARCC Level 4	PARCC Level 5	PARCC Level 4/5	Total
KRA Emerging Readiness	38.0	28.6	21.3	10.9	1.2	12.0	14,790
KRA Approaching Readiness	16.1	23.1	28.5	27.4	4.9	32.3	35,816
KRA Demonstrating Readiness	4.1	9.8	21.4	43.9	20.7	64.6	47,020
Total	13,347	17,103	23,418	32,064	11,694	43,758	97,626

Note. **Results are from the KRA 2014/PARCC 2018 and the KRA 2015/PARCC 2019 cohorts. PARCC levels were defined as: 1 = Did Not Yet Meet Expectations, 2 = Partially Met Expectations, 3 = Approached Expectations, 4 = Met Expectations, and 5 = Exceeded Expectations.**

3.2 KRA Reliability

In its simplest form, reliability measures the consistency of students' scores if the assessment were given multiple times or via multiple forms. Cronbach's alpha, which is a function of the number of items, the sum of all of the item variances, and the variance of the total scores, was used to evaluate reliability. Greater values of Cronbach's alpha (i.e., closer to 1) indicate that the items are closely related to one another and, additionally, that students score consistently across the items. The standard error of measurement is a function of the reliability measure (Cronbach's alpha) and is defined as the standard deviation of error scores for a student under repeated independent administrations with the same assessment (Allen & Yen, 1979).

Table 3.2.A summarizes the descriptive and reliability statistics, which fall within acceptable ranges, for the KRA in fall 2022. A complete summary of the descriptive and reliability statistics for all student groups is provided in the *Maryland KRA: Annual Technical Report* (2022–2023) (WestEd, 2023).

Table 3.2.A. Summary of Descriptive and Reliability Statistics for the KRA in Fall 2022

Domain	Mean	SD	Range	Cronbach's Alpha	SEM
Overall	266.32	15.52	202–298	0.95	3.33
Language and Literacy	264.96	17.30	202–298	0.88	5.93
Mathematics	264.86	18.16	202–298	0.83	7.39
Physical Well-Being and Motor Development	273.76	19.61	202–293	0.88	6.83
Social Foundations	271.63	21.60	202–298	0.93	5.88

Note. N = 59,893.

As described in Section 2.6 of this report, all early childhood educators who administer the KRA must complete training activities, including a simulator that models proper administration and scoring processes to support the reliability of item scores. Before any early childhood educator can administer the KRA, the educator must also pass a content assessment.

4 Differential Item Functioning (DIF) Analysis for the KRA

Differential item functioning (DIF) is a statistical concept that refers to a situation where there are differences in the probability of individuals from different groups (e.g., gender, ethnicity, language proficiency) responding correctly to a particular assessment item, despite having the same underlying level of ability or trait being measured. DIF is considered a relative term since it is always used when comparing one group of examinees to another on a given item (Holland & Wainer, 1993). DIF occurs when an item behaves differently for different groups of students who are assumed to have the same level of underlying ability (AERA et al., 2014). This discrepancy in item performance raises concerns about the fairness and validity of the assessment, as it suggests that the item may be biased toward one group. It is important to note that an item identified as behaving differently for different groups of students using a DIF analysis warrants further attention and review by a committee of experts, but it does not necessarily mean that the item is biased. It is the committee of experts that determines whether the item is indeed biased. In DIF analysis, it is customary to refer to the examinee group of interest, typically the marginal group, as the focal group. The group to which their performance is compared is referred to as the reference group (Bolt, 2002; Holland & Wainer, 1993). Typically, in any study, there could be multiple focal/reference pairs of groups for which DIF comparisons can be made.

4.1 Standardization Method for DIF

For this report, the DIF analysis was conducted using the standardization method for DIF (Dorans & Kulick, 1986), and the results were classified using the ETS DIF criteria (Zwick, 2012). The standardization method of studying DIF was developed for dichotomous items (with a 0/1 or incorrect/correct response), but this method is easily adapted to polytomous items that are scored on a rubric (Dorans & Holland, 1992). The standardization method takes the average difficulty value for each item at each score point, calculated for the focal groups and reference groups. The difference in the average difficulty values is weighted at each score point across all the score points to get the "standardized" difference at the item level.

The ETS criteria (Zwick, 2012) used to classify DIF results puts items into three categories: A (negligible or nonsignificant DIF), B (slight to moderate DIF), or C (moderate to large DIF).

Categories are further classified by a plus or minus (+/−) sign, indicating whether the item shows DIF in favor of the focal group or reference group, respectively. Therefore, an item with a C+ classification can be interpreted as possibly indicating significant DIF in favor of the focal group, while an item with a B− classification can be interpreted as possibly indicating moderate DIF in favor of the reference group. Zwick recommends using a minimum group size of 200 students in each group to have adequate statistical power to detect DIF.

The primary goal of DIF analysis is to find items that demonstrate significant DIF (i.e., C+ or C−) and flag them for review by a committee of experts. As previously mentioned, quantitative results from DIF analyses should be substantiated with a qualitative and critical review by a committee of subject matter experts. If the committee determines that these items are indeed biased, then these items should be excluded from the assessment, such that a fair assessment can be developed for members of all racial, ethnic, and gender subgroups (Angoff, 1993; Cole, 1993).

4.2 DIF Comparisons Conducted for the KRA

To evaluate the KRA items for DIF, the standardization method was employed on the KRA data from fall 2022. Student demographic groups evaluated for DIF (with focal and reference groups identified for each comparison) and the total number of students (*N*) in each group across all item comparisons are presented in Table 4.2.A.

Table 4.2.A. DIF Comparisons Conducted for the KRA in Fall 2022

Focal Group	Sample Size (*N*)	Reference Group	Sample Size (*N*)
Students who are economically disadvantaged	(*N* = 22,062)	Students who are not economically disadvantaged	(*N* = 37,827)
Female	(*N* = 29,101)	Male	(*N* = 30,785)
Black	(*N* = 18,200)	White	(*N* = 20,103)
Asian	(*N* = 3,888)	White	(*N* = 20,103)
Hispanic/Latino	(*N* = 13,694)	White	(*N* = 20,103)
English learners	(*N* = 9,608)	Not English learners	(*N* = 50,285)
Students with an IEP	(*N* = 5,969)	Students without an IEP	(*N* = 53,924)

4.3 DIF Results for the KRA

Table 4.3.A provides a summary of the DIF analysis by KRA domain for each pair of comparisons that were conducted (as listed in Table 4.2.A). Detailed results from the item-level DIF analyses for each item across all comparison groups are included in the Appendix.

Table 4.3.A. Total Number of Items With DIF by Domain for All Comparison Groups

Focal Group	Reference Group	Domain	A	B+	B−	C+	C−
Economically Disadvantaged	Not Economically Disadvantaged	LL	17				
		MA	13				
		PD	9				
		SF	11				
Female	Male	LL	17				
		MA	12		1		
		PD	5	2		2	
		SF	7	3		1	
Black	White	LL	16	1			
		MA	12		1		
		PD	9				
		SF	11				
Asian	White	LL	9		2	2	4
		MA	11	1	1		
		PD	8	1			
		SF	8	1	2		
Hispanic/Latino	White	LL	13		2		2
		MA	10	1	1		1
		PD	2	7			
		SF	7	2		2	
English Learner	Not English Learner	LL	13		1		3
		MA	10	1	1		1
		PD	1	6		2	
		SF	6		1	4	
Students with an IEP	Students without an IEP	LL	12	2	1	2	
		MA	9	4			
		PD	3		5		1
		SF	5		5		1

Note. LL = Language and Literacy, MA = Mathematics, PD = Physical Well-Being and Motor Development, SF = Social Foundations. Categories classified by a plus (+) sign indicate items that show DIF in favor of the focal group, and categories classified by a minus (−) sign indicate items that show DIF in favor of the reference group.

All items were classified as having negligible DIF (A) when compared by students' economic status. All other DIF comparisons resulted in one or more items with moderate to significant DIF.

For the female/male comparisons, 5 items (2 items in the Physical Well-Being and Motor Development domain and 3 items in the Social Foundations domain) were classified as B+, indicating that these items likely show moderate DIF in favor of female students. Moreover, 3 items (2 items in the Physical Well-Being and Motor Development domain and 1 item in the Social Foundations domain) were classified as C+, indicating significant DIF in favor of female students. One item (in the Mathematics domain) showed moderate DIF in favor of male students.

As shown in the Appendix, the Physical Well-Being and Motor Development items with moderate DIF in favor of female students (B+) have to do with either the use of small muscles (e.g., holding scissors and writing tools) or following safety rules. Similarly, the Social Foundations items with DIF in favor of female students have to do with managing the expression of feelings, thoughts, impulses, and behavior. However, given that all Physical Well-Being and Motor Development and Social Foundations items are observation-based items, these items should be specifically re-examined by subject matter experts to evaluate any implicit biases within the scoring rubrics to ensure that teachers are not evaluating boys more critically on these items.

For the Black/White comparisons, 1 Language and Literacy item was flagged as having moderate DIF in favor of Black students (B+), and 1 Mathematics item was flagged as having moderate DIF in favor of White students (B−). These items should also be further examined by subject matter experts for any implicit or explicit bias.

For the White/Asian comparisons, many items flagged for DIF were in the Language and Literacy domain. Six items in the Language and Literacy domain were classified as showing moderate to significant DIF in favor of White students (B− and C−), and two items were classified as showing significant DIF in favor of Asian students (C+). One Mathematics item was classified as showing moderate DIF in favor of Asian students (B+) and one Mathematics item was classified as showing moderate DIF in favor of White students (B−). One Physical Well-Being and Motor Development item was classified as showing moderate DIF in favor of Asian students (B+). And finally, two Social Foundations items were classified as showing moderate DIF in favor of White students (B−), and one Social Foundations item was classified as showing moderate DIF in favor of Asian students (B+).

As shown in the Appendix, some of the Language and Literacy items that show DIF in favor of White students (B− and C−) in the White/Asian comparisons assess higher-order language skills (e.g., naming verbs or using prepositions) or productive language skills (e.g., engaging in conversations or using words that reflect variety and complexity). It is possible that the development of these skills may be somewhat delayed for Asian students who are bilingual or

multilingual, or who are classified as English learners based on a home language survey, likely indicating that English may not be the primary language spoken at home by these children.

Further, as can be noted from the Appendix, some of these same items have been flagged for moderate or significant DIF (B− or C−) in the Hispanic/White and the English learner/not English learner comparisons, indicating that these items have been found to perform in favor of the White students or students who are not English learners. It is likely that these items are measuring language skills that are complex for either bilingual or multilingual students. Therefore, it is critical that these items be further evaluated by subject matter experts to evaluate their relevance to kindergarten readiness for bilingual or multilingual students.

Overall, for the Hispanic/White and English learner/not English learner comparisons, several items in the Physical Well-Being and Motor Development and the Social Foundations domains have been classified as having moderate to significant DIF (B+ or C+) in favor of the focal group (i.e., Hispanic/Latino students or English learner students). For the Hispanic/White comparison, 7 of the 9 Physical Well-Being and Motor Development items indicated moderate DIF in favor of the Hispanic/Latino students (B+), and 4 of the 11 Social Foundations items showed moderate to significant DIF in favor of the Hispanic/Latino students (B+ or C+). Similarly, for the English learner/not English learner comparison, 8 of the 9 Physical Well-Being and Motor Development items had moderate or significant DIF (B+ or C+) in favor of the English learner students, and 4 of the 11 Social Foundations items indicated significant DIF in favor of the English learner students (C+).

It might seem surprising that many items in the Physical Well-Being and Motor Development and Social Foundations domains are flagged for DIF in favor of the focal groups. However, in thinking carefully about the standardization method and how the comparisons are made, we should consider the fact that this method takes the average difficulty value for each item at each score point conditioned on the KRA total scores. Given that the KRA is considered a unidimensional construct—measuring a single underlying construct of kindergarten readiness—the item difficulties are conditioned on overall scores on the KRA. It is conceivable that bilingual and multilingual students might struggle with the Language and Literacy items (and to an extent the language-specific Mathematics items), but not have such a disadvantage in the observation-based Physical Well-Being and Motor Development and Social Foundations items.

Students of potentially higher underlying ability may have lower overall scores when all domain scores are combined due to the range restriction of these students' scores in the Language and Literacy and Mathematics domains. Therefore, it is possible that the Physical Well-Being and Motor Development and Social Foundations items do not have DIF in favor of the Hispanic/Latino and English learner students, but that Hispanic/Latino and English learner students of much higher underlying ability have lower overall scores due to the restricted range in which they can demonstrate their ability in the Language and Literacy and Mathematics domains.

To examine this hypothesis, we computed the average item difficulty by domain for students who are English learner and/or Hispanic/Latino, as shown in Table 4.3.B. The average item difficulty does not vary significantly by student group for the Physical Well-Being and Motor Development and Social Foundations domains. However, for the Language and Literacy and Mathematics domains, items appear to be significantly more difficult on average for students who are English learners, and particularly for Hispanic/Latino students who are English learners.

Table 4.3.B. Average Item Difficulty by Domain for Students who are Hispanic/Latino, Not Hispanic/Latino, and Students who are English Learner, Not English Learner

Mathematics (MA)		
	English Learner	Not English Learner
Hispanic/Latino	0.47	0.69
Not Hispanic/Latino	0.54	0.78

Language and Literacy (LL)		
	English Learner	Not English Learner
Hispanic/Latino	0.42	0.65
Not Hispanic/Latino	0.47	0.75

Social Foundations (SF)		
	English Learner	Not English Learner
Hispanic/Latino	0.68	0.79
Not Hispanic/Latino	0.69	0.84

Physical Well-Being and Motor Development (PD)		
	English Learner	Not English Learner
Hispanic/Latino	0.80	0.87
Not Hispanic/Latino	0.83	0.90

Note. **Item *p*-values are interpreted on a 0–1 scale, with higher values indicating easier items and lower values indicating harder items for those student groups.**

For comparisons of students with and without IEPs, 4 items in the Language and Literacy domain and 4 items in the Mathematics domain showed moderate to significant DIF in favor of the focal group (B+ or C+). Moreover, several items in the Physical Well-Being and Motor Development and Social Foundations domains showed moderate to significant DIF in favor of the reference group (B− or C−). It is possible that this may be expected for students who have an IEP because delays in physical and motor development are typically noted reasons for students' IEPs. Therefore, students with IEPs might have challenges completing the Physical Well-Being and Motor Development tasks compared to other students who have similar overall abilities. Further evaluation by subject matter experts for any implicit bias toward students with disabilities may be warranted.

5 Cultural Responsiveness of the Current KRA

In an increasingly diverse and interconnected world, the concept of cultural responsiveness has gained considerable prominence in education. Cultural responsiveness centers and values students' cultures and acknowledges that culture exerts a significant influence on all facets of social interactions, with particular significance in educational contexts, encompassing educators' instructional methods and students' growth and development (Ladson-Billings, 1995). Although a precise definition of culture presents challenges, cultural responsiveness recognizes that culture is a complex integration of explicit components, including behaviors, customs, language, and rituals, with implicit elements like beliefs, values, and social identity (Montenegro & Jankowski, 2017). Further, cognitive aspects of culture include how personal experiences shape knowledge, behavior, and communication (Walker et al., 2023). Understanding and implementing cultural responsiveness is not only a matter of academic significance but also reflects the values of equity, inclusion, belonging, and fairness.

In fall 2022, approximately 61,000 children began kindergarten in Maryland's public schools. More than half of the children are from historically underserved communities, as identified by race/ethnicity, English learner status, and economic disadvantage. Of this cohort of kindergarten students, approximately 30% identified as Black/African American, 34% identified as White, 23% identified as Hispanic/Latino, 7% identified as Asian, 6% identified as two or more races, and less than 1% identified as American Indian/Alaska Native or Native Hawaiian/Pacific Islander. Further, over 16% of entering kindergarteners are considered English learners, indicating that English is not the predominant language spoken at home, and nearly half of the children in the cohort were eligible for free or reduced-price lunch.

With the diverse identities of Maryland students who enter kindergarten, it is important to ensure that the KRA is culturally responsive and allows appropriate flexibility for all students across Maryland to demonstrate their school readiness. Therefore, in this section of the report, we first present a framework for evaluating assessments from a culturally responsive lens, adapted from the work by Walker et al. (2023), and subsequently use this framework to evaluate the purpose, design, development, administration, scoring, and reporting of the current KRA.

5.1 A Framework for Evaluating Assessments for Cultural Responsiveness

Culturally responsive assessment is defined as an evaluation that considers students' unique cultural identities and allows them the flexibility to demonstrate their mastery of a subject from their cultural perspective (Landl, 2021; Walker et al., 2023). Randall (2021) outlined an approach rooted in principles of equity and social justice, challenging the conventional practice of intentionally removing contextual elements from assessment items in pursuit of perceived neutrality. Culturally responsive assessments ensure that every student is represented throughout the design and development, promoting an equitable and fair assessment experience for all students (Walker et al., 2023).

A widely accepted heuristic for describing cultural responsiveness in education is the windows-and-mirrors metaphor. In this framework, windows represent opportunities for students to explore diverse perspectives and alternative worldviews, while mirrors reflect opportunities for students to identify with and see themselves represented in their educational experiences. Applying this metaphor to assessment design and development fosters equity and enhances student engagement, enabling learners to incorporate their own cultural backgrounds and references into the demonstration of their knowledge, skills, and behaviors (REL Pacific, 2021).

Walker et al. (2023) have outlined a set of five fundamental design principles for designing and evaluating culturally responsive assessments. They embed these five principles within all stages of assessment design, development, implementation, and use, and provide users with a framework to evaluate new and existing assessments for their cultural responsiveness. We found this framework useful in evaluating the current version of the KRA. The five fundamental design principles Walker et al. offer are as follows:

- **Shared power:** Utilizing input and experience from all concerned parties at all stages of the assessment process.

- **Engagement:** Fostering academic engagement and a sense of belonging in a way that reflects the students' identities and lived experiences.

- **High expectations:** Maintaining high expectations for all students that help negate biases that hinder student performance.

- **Flexibility:** Embracing flexibility to account for differences in culture, interests, and identities of all learners.

- **Asset-based:** Adopting an asset-based perspective that measures what students know and can do, and that disrupts traditional deficit narratives.

Further, Walker et al. (2023) provided a useful list of questions for users to consider when evaluating all stages of assessment development from a culturally responsive lens. Table 5.1.A includes an adaptation of this list, comprising principled questions to ask regarding the cultural

responsiveness of the KRA. Throughout the remainder of this section of the report, we address these questions in more detail for each stage of assessment.

Table 5.1.A. Principled Questions to Ask Regarding the Cultural Responsiveness of the KRA

Stage of Assessment	Principle	Questions to Ask
Purpose	Shared power; Engagement; Flexibility; High expectations; Asset-based	Who is eligible to take the KRA? Who is excluded? Who was included in the early conversation about the need for the KRA? How does the KRA provide students with opportunities to demonstrate school readiness? Who may be harmed? How desirable is the KRA to all involved parties?
Development	Shared power; Engagement; Flexibility; Asset-based	What structures allowed for co-design of the KRA and allowed for feedback from all involved parties at every stage of the KRA design and development? How are group differences in the definition of knowledge (i.e., school readiness) incorporated in the KRA?
Design	Shared power; Engagement; Asset-based	How does the KRA incorporate students' backgrounds? How does the KRA work to support students' strengths? How does the KRA allow students to show what they know?
Administration and Allowable Supports	Engagement; Flexibility; High expectations	How do KRA administration conditions help every student feel safe and welcome? How are individual differences among students accommodated?
Scoring and Reporting	Shared power; Flexibility; High expectations; Asset-based	What feedback is provided from the KRA? How does the KRA scoring model account for differences in mode of expression? How are views of involved parties incorporated into the KRA scoring model? How are students' rights protected? How are possible adverse consequences of KRA score use avoided? How does the use of KRA scores promote opportunity for all students? How are the KRA results communicated to all concerned parties?

Note. Adapted from Walker et al. (2023).

5.2 Evaluating the Cultural Responsiveness in the Purpose of the KRA

Since its inception in 2011, with the aid of Race to the Top – Early Learning Challenge grants from the U.S. Department of Education, the development of the KRA has been a collaborative effort between MSDE and its partners: the Ohio Department of Education, JHU CTE, WestEd, and a Technical Advisory Committee composed of early learning and assessment experts. Beginning in late 2013, with the aid of an Enhanced Assessment Grant from the U.S. Department of Education, MSDE, JHU CTE, WestEd, and numerous other states revised the KRA that was originally implemented in 2014 (i.e., KRA 1.5), resulting in the version of the KRA that has been administered in Maryland since 2018 (i.e., KRA 2.0). The primary objective of this collaborative effort was to establish a comprehensive KRA that effectively evaluates the essential domains of school readiness for all children entering kindergarten. Equally important was a commitment to ensure that the assessment would be accessible to young children from diverse backgrounds and with varying developmental needs (JHU CTE & WestEd, 2012).

The KRA has been a vital part of Maryland's state assessment program since its introduction in the fall of 2014. It began with census administration, meaning it was administered to all students, before transitioning to sample administration in some counties from 2016 to 2021. As of fall 2022, the KRA has returned to census administration across all counties in Maryland. Over the years, the KRA has become an integral component of Maryland's state assessment program. Recently, the KRA has become a key part of the AIB's comprehensive implementation plan, further supporting the Blueprint for Maryland's Future.

While numerous educators and community members were consulted throughout the development of the KRA, the purpose and design of the KRA was predominantly constructed by the state departments of education, their development partners, and the Technical Advisory Committee. Given that the KRA has been administered for nearly a decade, we recommend that MSDE bolster the previous approach taken in designing and developing the KRA and employ the "shared power" principle described in Section 5.1 to gather additional perspectives from a broader group of constituents and to substantiate the evidence that the KRA is meeting its intended claims and purposes for all involved parties in Maryland.

5.3 Evaluating the Cultural Responsiveness in the Development of the KRA

"Shared power" is critically important to ensure cultural responsiveness of the assessment for all concerned parties. This means that individuals from diverse constituencies across Maryland participate in the design and development of the KRA to ensure equitable outcomes for students from different backgrounds. Assessments developed by forming alliances with all impacted parties can additionally contribute to the validity argument of the assessment and the interpretation of results from those assessments (Walker et al., 2023).

The KRA's development followed an iterative process, detailed in Section 2 of this report, which involved extensive collaboration between MSDE staff, Maryland early childhood educators, and a technical advisory committee composed of national experts in early childhood education and development. Further, as detailed in Section 2 of this report, the KRA was developed following an iterative process that included cognitive interviews, a pilot, and two field tests. Each step of this process was designed to inform the subsequent step, providing critical evidence to support the validity and reliability of the KRA for its intended purpose. During each step of the development process, feedback was obtained from educators and students to ensure that these items were developmentally appropriate and accessible to all students entering kindergarten.

For example, educators and their students from various participating states were included in rounds of cognitive interviews to evaluate the KRA items for the extent to which students across these states were able to engage with the KRA item types. Further, all the items went through content and bias/sensitivity reviews prior to the pilot and field tests. These item-review committees included diverse panels of educators from Maryland and partnering states. The results from the cognitive interviews and the bias/sensitivity reviews were used to inform revisions to the KRA items that were used in the pilot, and results from the pilot were in turn used to inform additional revisions and modifications to items that were field tested, which were then used to inform the final operational items.

A large and representative sample of students and teachers from participating states were included in the pilot and field tests. During the pilot and field tests, teachers were asked to provide necessary accommodations for students with disabilities that were consistent with current state assessment policies and guidelines. Further, in a post-pilot survey, teachers were asked to answer specific questions about the accessibility of the KRA items, graphics, images, and manipulatives for English learners, and the need for allowances and accommodations for any of the items to make them more accessible to all students, including students with disabilities. Results from the pilot informed specific item enhancements focused on maximizing accessibility for English learners and students with disabilities. And, as detailed in Section 2.6.2, several supports are built in for diverse learners, and particularly for English learners and students with disabilities in the current version of the KRA.

However, as was noted in Section 2.4.1 of this report, students who are English learners or students with severe cognitive disabilities were not included in the cognitive interviews. Therefore, specific challenges that students from these groups may have in engaging with the KRA items and item types have not been explicitly observed and documented. There was also no intentional inclusion of students from historically marginalized groups in these cognitive interviews. Therefore, in revising the KRA to be more culturally responsive, students who are English learners, students with disabilities, and students from historically marginalized groups should be explicitly and intentionally included in the pilot studies to observe and document how students from these groups are able to engage with the KRA.

5.4 Evaluating the Cultural Responsiveness in the Design and Content of the KRA

Assessments that are designed to be engaging by infusing culturally relevant and context-rich content could be beneficial to all students. The "engagement" and "asset-based" principles proposed by Walker et al. (2023) are most relevant when evaluating an assessment's design and content. To evaluate the KRA based on this principle, key questions to answer include the extent to which the KRA blueprint and items incorporate diverse student backgrounds, engage students' strengths, and allow students to show what they know and are able to do (see Table 5.1.A). Therefore, two researchers independently applied a culturally responsive lens to evaluate the KRA blueprint and items in order to specifically answer the principled questions outlined in Table 5.1.A. Their review of the cultural responsiveness of the KRA design (blueprint and item types) and content (items and forms) are detailed in the subsequent sections.

5.4.1 Evaluating the KRA Blueprint

As summarized in Section 2.3 of this report, the current KRA blueprint includes four domains (Social Foundations, Language and Literacy, Mathematics, and Physical Well-Being and Motor Development) that allow children to demonstrate different areas of strengths. The distribution of item score points across the four domains is relatively balanced, with each domain representing at least 19% of the points. The Language and Literacy domain includes the most points (35% of the total points); Social Foundations and Mathematics each represent 23% of the total points; and Physical Well-Being and Motor Development represents 19% of the total points. The blueprint also indicates that approximately half of the items (26) are selected-response items and performance-task items and approximately half of the items (24) are observation-based items, allowing students to demonstrate their strengths and abilities in various ways.

A more detailed review of the blueprint and item specifications indicate that the following topics are included in the KRA: social-emotional development; executive functions; reading, writing, language, speaking, and listening skills; counting and cardinality; number sense and operations; measurement; geometry; physical education; and basic health skills. While the inclusion of these various skills allows students to demonstrate their varied strengths and all of these skills seem relevant to children who are entering kindergarten, the KRA could benefit from additional evaluation of whether all of these skills are relevant to children from various cultural backgrounds, English learners, students with disabilities, and neurodivergent children.

In addition, the current blueprint does not explicitly denote culture. As previously mentioned, the blueprint includes the specific domains and learning standards for which the KRA items align; however, the blueprint could be augmented to also include criteria of cultural responsiveness. These criteria would include categories such as race/ethnicity, gender, disability, and geographic location (i.e., urban, suburban, rural). Additionally, each item could

be labeled with these additional categories to identify each item's cultural context. Such a process offers the opportunity to infuse cultural responsiveness by including examples, scenarios, characters, pictures, and contexts from diverse cultures represented within the Maryland student population throughout the entire assessment. Subsequently, a review of the KRA items could be conducted to ensure that they are representative of children from various cultural backgrounds and students with divergent abilities.

5.4.2 Evaluating the KRA Items

The current design of the KRA allows for cultural responsiveness in various item types, particularly in performance tasks, where a student's cultural or linguistic background can be considered. Further, the KRA includes many images that are typical and necessary for kindergarten assessments, such as animals, trees, balls, blocks, planes, and eating utensils, and seem to cross over multiple cultures. The illustrations of people seem somewhat diverse but also racially ambiguous. There are a group of items that ask about details in a story. The story's theme is about a boy who goes shopping with his grandma. Observational items include a variety of tasks and scenarios, including familiarity with print, verbal communication, writing, social-emotional and behavioral skills, receptive language skills, and fine and gross motor skills. These observational items are observed during regular school activities, allowing students various ways to demonstrate strengths. Further, the observational items are rated by a rubric that utilizes positive and "asset-based" language, specifically allowing teachers to document skills and behaviors that are evident, in progress, or not *yet* evident.

While these items are likely relevant to most children, there are opportunities to include representations of experiences and objects that students from diverse cultural backgrounds might encounter on a regular basis. Moreover, the items could also be evaluated to ensure that they provide students with an opportunity to demonstrate their strengths. Some items on the KRA encourage teachers to prompt the child or gather more information if the child's initial response isn't detailed enough. However, such guidance and flexibility are not provided for most other items to identify and highlight student strengths, and this flexibility may be beneficial across more items on the KRA. In addition, intentional consideration could be given to enhancing the diversity of the images and illustrations included in the KRA by reflecting children's race, ethnicity, ability, and geographic differences (e.g., suburban park vs. urban playground; small corner store vs. large grocery store; kitchen/home that looks culturally inviting vs. the somewhat "sterile" environment illustrated in the current items).

Observational-rubric items should be sensitive to how children from different cultural backgrounds may differ in their demonstration of the skills and behaviors. Several cultures might prioritize certain behaviors or skills in different ways. While the current observational-rubric items use asset-based language, additional guidance should be provided to teachers to ensure that they consider the diversity of students' experiences in their observations of students from diverse backgrounds. For example, in the section for social-emotional learning,

examples and scenarios from various cultural backgrounds could be included to help teachers understand that emotions can be expressed differently across various cultures. Without additional guidance to understand how students from diverse backgrounds have been taught to respond to situations in culturally appropriate ways, teachers may extend their own worldviews in their observations of students from diverse cultures.

In revisiting the KRA for its cultural responsiveness, it is important to establish an advisory panel that includes representatives and community members from diverse constituencies, specifically those that represent students from historically marginalized groups, English learners, and students with disabilities. This advisory panel could evaluate the current KRA blueprint and items from a culturally responsive lens to ensure that the KRA items and the item types are accessible to all students entering kindergarten.

5.5 Evaluating the Cultural Responsiveness in the Administration of the KRA

One aspect of creating culturally responsive guidelines for administration is to allow for "flexibility" in access, presentation, and accommodations for multiple response formats. This means that the needs and preferences of all concerned parties should be considered to promote a sense of autonomy, competence, and belonging among the students being assessed. To evaluate the KRA for its cultural responsiveness in its administration, particular attention should be paid to the administration conditions (i.e., do all students feel safe and welcome?) and the extent to which the KRA provides accommodations for individual differences among the students being assessed.

The KRA is administered in classroom settings by classroom teachers who have gained some familiarity with the students since the start of the school year. Several items are observation-based, and extensive professional development is offered to teachers who administer the KRA to ensure that all students have a consistent experience. Teacher administration of the KRA currently includes differentiable allowances based on universal design, and clearly provides flexibility in directions, item presentation, student response, setting, and scheduling. All the accommodations provided are more fully summarized in the *Guidelines on Allowable Supports for the Kindergarten Readiness Assessment*.[9] Overall, multiple means of engagement have been built in to the KRA, particularly for students who are English learners and students with disabilities. For example, supports are available for English learners and students with disabilities by accepting answers in the child's native language and a defined ASL protocol. Students can also respond with gestures in lieu of verbal responses. These allowances provide opportunities for students to respond in a "manner in which they are most comfortable and confident" (Walker et al., 2023, p. 8).

---

[9] These guidelines are available at: https://pd.kready.org/data/ck/sites/116/files/MD_KRA%20Guideline%20final%20rv.pdf

In ensuring that the KRA provides flexibility that considers the interests and needs of all impacted members of the community, Walker et al. (2023) recommend expanding on the principles of universal design for learning (UDL) (CAST, 2018). UDL guidelines currently include three overarching principles to allow for multiple means of presenting content, multiple ways for students to respond, and multiple means of engagement. In addition, cultural diversity could be added to these guidelines as a key area of focus, paying particular attention to the diverse funds of knowledge that different students bring to the administration conditions. Paying attention to diversity of experiences is especially important for very young students, whose main point of reference is what is happening in their homes and families. Therefore, in expanding the flexibility afforded within the KRA to diverse learners, we recommend that observation-based items and rubrics allow for multiple ways of demonstrating skills, knowledge, and behaviors so that the development of students from diverse backgrounds is taken into consideration. Such an intentional consideration of flexibility may be accomplished by internalizing the "shared power" principle and engaging the advisory panel from diverse constituencies during the KRA revision process.

5.6 Evaluating the Cultural Responsiveness in the Scoring and Reporting of the KRA

It is important that assessments are scored and reported in ways that ensure high expectations for all students. Therefore, in evaluating the scoring and reporting of the KRA, it is important to verify if the feedback provided is designed to motivate all students; the scoring model accounts for differences in modes of expression; scores are reported in a way that is interpretable and useful to all relevant groups; and steps are taken to ensure that the scores are used appropriately and do not lead to any adverse consequences for children taking the KRA.

Section 2.5 of this report discusses scoring and reporting efforts undertaken for the current KRA. As detailed in Section 3 of this report, and further in the development and technical report of the KRA (WestEd, 2018), concerted efforts were taken to ensure the validity and reliability of the KRA scores for all students. Furthermore, various types of score reports have been designed for educators and parents and made available both through the KReady system and shared with parents as PDF documents. Caution about the appropriate uses and interpretations of the KRA scores have been laid out in the Individual Student Report (ISR) for parents and in the educator reports. In addition, professional development opportunities are provided for teachers in administering, scoring, and interpreting the KRA reports to ensure that the KRA results are appropriately used to provide necessary support to students where needed and does not lead to adverse consequences for any student.

In addition, the content of the current KRA reports could be enhanced for cultural responsiveness. This can be achieved by starting with the current reports and actively seeking input from a diverse range of relevant groups, including teachers, parents, and community members from heterogeneous groups. Their valuable perspectives can help ensure that the

results reported are interpretable and useful to all relevant groups (AERA et al., 2014). Therefore, the KRA score reports should be re-evaluated using a user-centric and iterative multistep framework (Kannan, 2023) to ensure that diverse groups of intended users who represent each constituency are involved in the co-design and development of the reports (Kannan et al., 2018). The review of the score reports could start with an intentional "needs assessment" with all concerned parties to understand gaps in interpretation, appropriate use, and usefulness of the current score reports. To be most useful to the intended users, the score report design and development process should occur in conjunction with the assessment review and revision process (i.e., when engaging a diverse advisory panel) to ensure that the needs of various intended user groups can be met in the claims being made by the assessment and vice versa.

6 Conclusions and Next Steps

The main purpose of this report was to evaluate the current version of the KRA for any racial, cultural, or linguistic biases in the KRA. To this end, a thorough review of the design and development of the KRA, accommodations and supports for students with disabilities and students who are English learners, scoring and reporting procedures, evidence of reliability and validity, and an evaluation of the cultural responsiveness of the KRA was undertaken. In addition, the results from the administration of the KRA in fall 2022 were evaluated for differential item functioning (DIF) across various groups of Maryland kindergarten students by gender, race/ethnicity, English language proficiency, and socioeconomic status.

Overall, the results from the DIF analysis and the review of the cultural responsiveness of the KRA indicate that some aspects of the current KRA, and particularly some items, need to be reviewed and evaluated by subject matter experts for any implicit bias toward students from various racial, ethnic, and linguistic backgrounds. For example, the DIF analysis showed that some Language and Literacy items that measured applied language skills (e.g., naming nouns and verbs, using prepositions) or productive language skills (e.g., engaging in conversations and using words that reflect variety and complexity) were challenging for Asian students, Hispanic/Latino students, and students who are English learners. We recommend that these items be re-evaluated by subject matter experts in item-review meetings, where educators can decide whether the items need to be revised to best serve the diverse needs and cultural backgrounds of Maryland's kindergarten students. These reviews should particularly consider the relevance of some of these skills (e.g., applied language skills) for the purpose of assessing children's readiness for kindergarten.

Further, the DIF analysis and the follow-up evaluation of item difficulties for Hispanic/Latino students and students who are English learners showed that there could be a significant disadvantage for students who are bilingual or multilingual learners, specifically in their ability to access some items within the Language and Literacy and Mathematics domains. Therefore, we recommend that items for these domains be evaluated in item-review meetings with subject matter experts to identify any potential implicit or explicit bias in these items.

We also recommend the development of a KRA section that assesses bilingual or multilingual children's skills in their home or preferred language to ensure that more actionable information can be provided to educators and families to better support these students' learning needs.

Presumably, children who cannot demonstrate readiness in mathematics and language and literacy when being assessed in English might be able to demonstrate these skills in their home or preferred language. There is also strong evidence to support that assessing bilingual or multilingual students in their home language <u>and</u> in English provides the most accurate estimates of their overall ability level (Pitoniak et al., 2009; Durán et al., 2022). If teachers and families can use KRA results to better understand students' abilities in English and their home or preferred language, then better instructional decisions can be made to support these children's learning needs. Children who speak Spanish at home represent about 23% of Maryland's public-school students; therefore, we strongly recommend that MSDE develop KRA sections to be administered in Spanish, in addition to the current version in English. While this will require that MSDE establish additional KRA policies to support decisions about which children to administer the Spanish-language version to, doing so would provide MSDE and Maryland educators with a more accurate picture of children's learning and development upon entry to kindergarten.

Students have a range of lived experiences, and their assets at the intersection of multiple characteristics (e.g., gender, race, ethnicity, language, disability status, religion, personal interests and values) should be identified and incorporated in reviewing and revising the KRA, including its scoring procedures and observational rubrics, to ensure that the KRA measures what students know and are able to do and disrupts traditional deficit narratives about diverse learners. The cultural diversity within the state's student body should be appropriately reflected in the assessments to ensure that they align with the needs, values, preferences, and diverse cultural backgrounds and experiences of the students.

One approach to ensuring the cultural responsiveness of the KRA is to emphasize the "shared power" principle (Walker et al., 2023). This could be accomplished by convening an advisory panel that includes representatives from MSDE; representatives from various school districts, jurisdictions, and governmental agencies across Maryland; and community members from diverse constituencies, specifically those that represent students from historically marginalized groups. This advisory panel could then be engaged in all phases of the KRA's review and revision process, starting with an evaluation of the current KRA blueprint, items, and reports. Such an intentionally forged partnership that focuses on the experiences of diverse interest groups and ensures all voices are heard throughout the review and revision process is key to achieving cultural validity of the KRA.

To ensure that the KRA provides "flexibility" that considers the interests and needs of all impacted members of the community, Walker et al. (2023) recommend adding cultural diversity to the principles of universal design, paying close attention to the diversity of experiences and funds of knowledge that diverse learners bring from their homes and families into their kindergarten year. Therefore, in reviewing and revising the KRA administration procedures, the developmental diversity of students should be intentionally considered when engaging the advisory panel.

Finally, the current KRA score reports should be re-evaluated for cultural responsiveness by actively seeking input from a diverse range of interest groups, including teachers, parents, and community members, to ensure that the reports are interpretable and useful to all relevant parties. The focus of the score report review should be to understand gaps in interpretation and to ensure that the needs of various intended user groups can be met in the claims being made by the KRA.

In summary, we recommend the following next steps in reviewing and revising the KRA:

1. Convene an advisory panel of experts and educators from diverse constituencies across Maryland (e.g., MSDE staff, Maryland educators, staff from state universities, representatives of state/local government agencies, and community members) to review the KRA blueprint, items, reports, and administration processes and policies to ensure that the KRA is responsive to the numerous cultures and backgrounds of Maryland's kindergarten students.

2. Conduct item-review meetings with subject-matter experts and educators to specifically review the items flagged in the DIF analysis and to determine whether these items require revisions to ensure that they are free from any potential bias. Revise items, if necessary.

3. Develop an additional section of the KRA that evaluates Spanish language and literacy proficiency for students whose home or preferred language is Spanish. This new Spanish-language section should be piloted and field tested prior to implementation within the operational KRA. Further, this additional section should be administered in addition to the domains currently included in the KRA. This would also require that MSDE establish policies to support the administration of this additional section.

4. Review and revise the score reports to ensure that they are interpreted and used appropriately by all intended users and constituents, especially teachers and families.

References

Allen, M. J., & Yen, W. M. (1979). *Introduction to measurement theory*. Brooks/Cole.

American Educational Research Association (AERA), American Psychological Association (APA), & National Council on Measurement in Education (NCME). (2014). *Standards for educational and psychological testing*. American Educational Research Association.

Angoff, W. H. (1993). Perspectives on differential item functioning methodology. In P. W. Holland & H. Wainer (Eds.), *Differential item functioning* (pp. 3–23). Lawrence Erlbaum Associates.

Bolt, D. M. (2002). A Monte Carlo comparison of parametric and nonparametric polytomous DIF detection methods. *Applied Measurement in Education*, *15*(2), 113–141. https://doi.org/10.1207/S15324818AME1502_01

CAST. (2018). *The universal design for learning guidelines* (version 2.2). http://udlguidelines.cast.org

Cole, N. S. (1993). History and development of differential item functioning. In P. W. Holland & H. Wainer (Eds.), *Differential item functioning* (pp. 25–33). Lawrence Erlbaum Associates.

Dorans, N. J., and Holland, P. W. (1992). *DIF detection and description: Mantel-Haenszel and standardization*. ETS Research Report Series. https://doi.org/10.1002/j.2333-8504.1992.tb01440.x

Dorans, N. J., & Kulick, E. (1986). Demonstrating the utility of the standardization approach to assessing unexpected differential item performance on the Scholastic Aptitude Test. *Journal of Educational Measurement*, *23*(4), 355–368. https://doi.org/10.1111/j.1745-3984.1986.tb00255.x

Dragoset, L., Baxter, C., Dotter, D., & Walsh, E. (2019). *Measuring school performance for early elementary grades in Maryland*. U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Mid-Atlantic.

Durán, L., Basaraba, D., Baker, D., & O'Brien, G. (2022). *The importance of dual language assessment in early literacy*. Amplify.

Holland, P. W., & Wainer, H. (Eds.). (1993). *Differential item functioning* (1st ed.). Routledge. https://doi.org/10.4324/9780203357811

Johns Hopkins University Center for Technology in Education & WestEd. (2012). *A conceptual design document: The early childhood comprehensive assessment system development, implementation, and evaluation report*.

Kannan, P. (2023). Score reporting: Design and evaluation methods informed by research. In R. J. Tierney, F. Rizvi, and K. Ercikan (Eds.), *International encyclopedia of education* (4th ed., pp. 217–229). Elsevier.

Kannan, P., Zapata-Rivera, D., & Leibowitz, E. A. (2018). Interpretation of score reports by diverse subgroups of parents. *Educational Assessment*, *23*, 173–194. https://doi.org/10.1080/10627197.2018.1477584

Ladson-Billings, G. (1995). Toward a theory of culturally relevant pedagogy. *American Educational Research Journal*, *32*(3), 465–491.

Landl, E. (2021, October). *Culturally responsive assessment and balanced assessment systems* [Webinar]. REL Pacific Webinar. https://www.youtube.com/watch?v=rT3tHRyCMhA

Montenegro, E., & Jankowski, N. A. (2017). *Equity and assessment: Moving towards culturally responsive assessment* (Occasional Paper No. 29). National Institute for Learning Outcomes Assessment. https://www.learningoutcomesassessment.org/wp-content/uploads/2019/02/OccasionalPaper29.pdf

Paris, D., & Alim, H. S. (2014). What are we seeking to sustain through culturally sustaining pedagogy? A loving critique forward. *Harvard Educational Review*, *84*(1), 85–100.

Pitoniak, M., Young, J. W., Martiniello, M., King, T. C., Buteux, A., & Ginsburg, M. (2009). *Guidelines for the assessment of English learners*. Educational Testing Service.

Randall, J. (2021). "Color-neutral" is not a thing: Redefining construct definition and representation through a justice-oriented critical antiracist lens. *Educational Measurement: Issues and Practice*, *40*(4), 82–90. https://doi.org/10.1111/emip.12429

Regional Educational Laboratory Pacific. (2021). *Culturally responsive assessment: Goals, challenges, and implications*. [Infographic.] https://ies.ed.gov/ncee/rel/infographics/pdf/REL_PA_Culturally_Responsive_Assessment.pdf

Sato, E., Rabinowitz, S., Gallagher, C., & Huang, C. W. (2010). *Accommodations for English language learner students: The effect of linguistic modification of math test item sets*. (NCEE 2009–4079). National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Walker, M. E., Olivera-Aguilar, M., Lehman, B., Laitusis, C., Guzman-Orth, D., & Gholson, M. (2023). *Culturally responsive assessment: Provisional principles* (Research Report No. RR-23-11). Educational Testing Service. https://doi.org/10.1002/ets2.12374

WestEd. (2014). *Ready for kindergarten: Kindergarten readiness assessment technical report*.

WestEd. (2023). *Maryland Kindergarten Readiness Assessment (KRA): Annual technical report (2022–2023)*.

WestEd. (2018). *Kindergarten readiness assessment 2.0 development and technical report*.

Wise, S. L., & Smith, L. F. (2016). The validity of assessment when students don't give good effort. In G. T. L. Brown & L. R. Harris (Eds.), *Handbook of human and social conditions in assessment* (pp. 204–220). Routledge.

Zwick, R. (2012). *A review of ETS differential item functioning assessment procedures: Flagging rules, minimum sample size requirements, and criterion refinement* (ETS Research Report No. RR-12-08). Educational Testing Service. https://doi.org/10.1002/j.2333-8504.2012.tb02290.x

Appendix: Differential Item Functioning (DIF) Statistics by KRA Item

Domain	Item	Item Description	DIF Comparisons						
			ED/ Not ED	Female/ Male	Black/ White	Asian/ White	Hispanic/ White	EL/ Not EL	IEP/ No IEP
LL	A322	Identify beginning sound	A	A	A	A	A	A	A
LL	A323	Identify beginning sound	A	A	A	A	A	A	A
LL	A326	Identify rhyming words	A	A	A	A	A	A	A
LL	A327	Identify rhyming words	A	A	A	B-	A	A	A
LL	A331	Make letter sounds	A	A	A	A	A	A	B+
LL	A333	Name letters (uppercase)	A	A	A	C+	A	A	C+
LL	A334	Name letters (lowercase)	A	A	B+	C+	A	A	C+
LL	A336	Distinguish words from letters	A	A	A	A	A	A	B+
LL	A343	Name nouns and verbs	A	A	A	C-	C-	C-	A
LL	A345	Use prepositions	A	A	A	C-	C-	C-	A
LL	B317	Answer question about story detail	A	A	A	C-	B-	B-	A
LL	B320	Identify story sequence	A	A	A	A	A	A	A
LL	B385	Answer question about story detail	A	A	A	A	A	A	A
LL	R340	Engage in conversations	A	A	A	C-	A	A	B-
LL	R341	Write first name	A	A	A	A	A	A	A
LL	R346	Use words that reflect variety and complexity	A	A	A	B-	B-	C-	A
LL	R401	Demonstrate how print is read	A	A	A	A	A	A	A

Domain	Item	Item Description	DIF Comparisons						
			ED/Not ED	Female/Male	Black/White	Asian/White	Hispanic/White	EL/Not EL	IEP/No IEP
MA	A348	Count to 20	A	A	A	A	A	A	A
MA	A349	Tell how many	A	A	A	A	A	A	A
MA	A351	Name numerals	A	B-	A	B+	A	A	B+
MA	A354	Pair numerals with sets	A	A	A	A	A	A	B+
MA	A355	Solve simple addition problem	A	A	A	A	A	A	A
MA	A358	Determine amount needed to complete a set	A	A	A	A	A	A	A
MA	A359	Sort by one attribute	A	A	A	A	A	A	A
MA	A362	Identify set that is "less than"	A	A	A	A	A	A	A
MA	A363	Identify set that has "the same number as"	A	A	A	A	A	A	A
MA	A366	Compare height	A	A	A	B-	C-	B-	A
MA	A367	Order objects by size	A	A	B-	A	A	A	A
MA	A370	Match shapes	A	A	A	A	B+	B+	B+
MA	A372	Name shapes	A	A	A	A	B-	C-	B+
PD	R373	Move through a crowded environment	A	B+	A	A	B+	C+	B-
PD	R378	Hold scissors	A	B+	A	A	A	B+	A
PD	R380	Hold writing tool	A	C+	A	B+	B+	C+	A
PD	R381	Follow basic safety rules	A	C+	A	A	B+	B+	B-
PD	R382	Explain ways adults help to keep people safe	A	A	A	A	A	A	B-
PD	R383	Complete personal care tasks	A	A	A	A	B+	B+	C-
PD	R384	Follow basic health practices	A	A	A	A	B+	B+	B-
PD	R397	Run, jump, and/or hop	A	A	A	A	B+	B+	B-
PD	R398	Bend, stretch, and/or twist	A	A	A	A	B+	B+	A

Domain	Item	Item Description	DIF Comparisons						
			ED/ Not ED	Female/ Male	Black/ White	Asian/ White	Hispanic/ White	EL/ Not EL	IEP/ No IEP
SF	R300	Express own emotions	A	A	A	A	A	A	B-
SF	R302	Ask familiar adults for help	A	A	A	B-	A	A	A
SF	R306	Wait to take turns	A	B+	A	A	C+	C+	A
SF	R307	Focus on activities	A	B+	A	A	B+	C+	A
SF	R309	Follow multi-step directions	A	A	A	A	A	A	B-
SF	R310	Solve problems	A	A	A	A	A	A	B-
SF	R312	Express a desire to learn	A	A	A	A	A	A	A
SF	R313	Engage with peers in pretend play	A	A	A	A	A	A	C-
SF	R314	Share materials with peers	A	B+	A	A	B+	C+	B-
SF	R315	Explain why rules are needed	A	A	A	B-	A	B-	B-
SF	R400	Control impulses and behavior	A	C+	A	B+	C+	C+	A

Note. **In each comparison, focal group is listed first followed by reference group: ED = Economically Disadvantaged; EL = English Learner; IEP = Students with Individualized Educational Plan; LL = Language and Literacy, MA = Mathematics, PD = Physical Well-Being and Motor Development, SF = Social Foundations**