



**SMU**

ANNETTE CALDWELL SIMMONS

SCHOOL OF EDUCATION & HUMAN DEVELOPMENT

RESEARCH IN MATHEMATICS EDUCATION

# **Imagination Station (Istation): Updates to Universal Screener Instrument Development for Grades 2-8**

RESEARCH IN  
MATHEMATICS  
EDUCATION

# **Imagination Station (Istation): Updates to Universal Screener Instrument Development for Grades 2-8**

Cassandra Hatfield • Brea Ratliff • Saller Axel • Deni Basaraba • Leanne Ketterlin-Geller  
Southern Methodist University

**Spring 2015**

Published by:

Southern Methodist University  
Department of Education Policy & Leadership  
Simmons School of Education & Human Development  
PO Box 750114  
Dallas, TX 75275-0114  
Contact information: [rme@smu.edu](mailto:rme@smu.edu)

This research was supported by Imagination Station Inc. Opinions expressed herein do not necessarily reflect those of Imagination Station Inc. or individuals within.

Copyright © 2015. Southern Methodist University. All rights reserved. This publication, or parts thereof, may not be used or reproduced in any manner without written permission.

SMU will not discriminate in any employment practice, education program or educational activity on the basis of race, color, religion, national origin, sex, age, disability or veteran status. This document is available in alternative formats upon request.

## **Abstract**

---

In this technical report, we describe the development of an updated item bank for the Grade 2 through Grade 8 formative assessment for Imagination Station (Istation). The formative assessment item bank will be used to deliver computerized-adaptive universal screeners to support teachers' instructional decision-making. State and national mathematics content standards inform the mathematics topics underlying the items. In this technical report, we describe the process used to identify and sample the mathematics content and levels of cognitive complexity assessed in the item bank. Next, we describe the item writing procedures and provided the qualifications for the item writers. Finally, we describe how the external item review process and outcomes contribute to content-related evidence for validity.

## Table of Contents

---

Introduction	1
Construct Definition	1
Levels of Cognitive Engagement	2
Item Writing	3
Item Specifications	3
Item Writing Procedures	3
Content-Related Evidence for Validity	5
Mathematics Education Expert Review	5
Mathematics Teacher Review	12
Conclusion	21
References	22
Appendix A: Grade 2 External Review Data	23
Appendix B: Grade 3 External Review Data	25
Appendix C: Grade 4 External Review Data	27
Appendix D: Grade 5 External Review Data	29
Appendix E: Grade 6 External Review Data	31
Appendix F: Grade 7 External Review Data	33
Appendix G: Grade 8 External Review Data	35
Appendix H: Item Writer Biographies	37
Appendix I: Mathematics Teacher Reviewer Biographies	39
Appendix J: Mathematics Education Expert Reviewer Biographies	42
Appendix K – State Content Standards Referent Sources	44

# **Imagination Station (Istation): Updates to Universal Screener Instrument Development for Grades 2-8**

## **Introduction**

The purpose of the Grade 2-8 formative assessment item bank for Imagination Station Indicators of Progress (ISIP) Mathematics universal screeners is to support teachers' instructional decision-making. The formative assessment item bank is used to deliver computerized-adaptive universal screeners to identify students' understanding of fundamental mathematics skills and grade-level standards. By administering this assessment system, teachers and administrators can use the results to answer these questions: (a) are students at risk of failure in grade-level mathematics, and (b) what is the degree of intensity of instructional support students need to be successful in grade-level mathematics? Multiple administrations of the universal screener (i.e., fall, winter, and spring each year) provide teachers with meaningful information about students' learning over time, and support instructional decision-making over the course of the academic year. The universal screener is designed for administration to all students receiving grade-level instruction.

The purpose of this technical report is to describe the development of an update to the existing formative assessment item bank. This description includes: (a) the process used to identify and sample the mathematics content assessed in the item bank, (b) the item writing process, and (c) the external review process and results. The test development steps used to create the formative assessment item bank represent best practices in test development and align with the test standards published by the American Educational Research Association (AERA), American Psychological Association (APA), and National Council on Measurement in Education (NCME) (1999).

## **Construct Definition**

The assessed construct consists of mathematics content and levels of cognitive engagement (National Research Council [NRC], 2001) elicited by each item. The Common Core State Standards for Mathematics (CCSS-M), and state content standards from Texas and Virginia (See Appendix K) provided the mathematics content framework for the formative assessment item bank. Revisions and additions to the formative assessment item bank were needed for two reasons: (1) many states (approximately 40) have adopted the CCSS-M and are no longer using the Curriculum Focal Points published by the National Council of Teachers of Mathematics (NCTM, 2006) as the basis for their content standards, and (2) the state content standards in Texas were revised in 2012 and implemented starting the 2013-2014 school year.

We began the process of updating the assessment blueprint for each grade level by engaging in discussions with experienced mathematics educators and mathematics researchers. First, we aligned the Texas state content standards (Texas Education Agency, 2012) with the CCSS-M and Virginia Standards for Learning (Virginia Department of Education, 2009). During this step we

focused on in-grade alignment; for example, ensuring that a Grade 3 CCSS-M content standard was aligned to a Grade 3 Texas standard. Out of grade alignments were not made. Next, we cross-referenced the codes from the previously developed blueprint with the newly created blueprint. Some standard codes remained unchanged while other content standards were coded to a new grade level to reflect the new content standards. Finally, we identified each content standard as *assessable or not assessable*. A small number of standards categorized as *not assessable* consisted of skills that could not be assessed using an online multiple-choice format (e.g., fluency with operations.) Determinations were made based on expert opinion and were verified by at least two members of the Research in Mathematics Education (RME) staff, all of whom have extensive mathematics education experience.

## Levels of Cognitive Engagement

---

The cognitive engagement dimension of the construct refers to the level of cognitive processing through which students are expected to engage with the content. The formative assessment item bank uses the taxonomy of cognitive engagement in mathematics published by the NRC (2001). The taxonomy consists of five interdependent strands that promote mathematical proficiency: (a) conceptual understanding, (b) procedural fluency, (c) strategic competence, (d) adaptive reasoning, and (e) productive disposition. Items in the formative assessment item bank assess student understanding of the content at four levels of cognitive engagement. Productive disposition was not assessed because it refers to a student's overall perception of mathematics, and personal belief in one's own efficacy in solving problems. A brief description of each assessed level follows:

- *Conceptual understanding* pertains to the functional grasp of mathematics that a student applies to concepts, operations, and relations. It involves being able to logically organize one's knowledge to integrate and understand concepts as part of a coherent whole.
- *Procedural fluency* pertains to students' ability to accurately and appropriately carry out skills, including being able to select efficient and flexible approaches.
- *Strategic competence* involves one's ability to formulate a problem in mathematical terms, to represent it strategically (verbally, symbolically, graphically, or numerically), as well as to solve it effectively. It is similar to problem solving and problem formation.
- *Adaptive reasoning* involves the student's capacity to think logically about a problem, which requires reflecting on various approaches to solve a problem and deductively selecting an approach. Students who are able to do this are also able to rationalize and justify their strategy.

Each content standard was examined to determine if it could be assessed at the four levels of cognitive engagement. Conceptual understanding and procedural fluency were oversampled to represent a balanced approach to mathematics proficiency as described by the CCSS-M. Easy, medium, and difficult items were written for each standard across the four levels of cognitive engagement. The level of difficulty of each item is a relative description that is subject to change with empirical analyses. The number of items written per content standard was based on these criteria: (a) the number of items previously written to the standard, (b) the number of states that

had alignment to the standard. The content sampling matrix is presented in Table 1 of Appendices A-G, with each appendix representing a grade level (e.g., Appendix A is Grade 2).

## **Item Writing**

### **Item Specifications**

---

Approximately 300 items were written per grade. Multiple-choice items were created for efficiency in the computer delivery system. Each item had three distractors and one correct answer. Items were written for dichotomous scoring as either correct or incorrect. The distractors represent plausible misconceptions or errors in conceptual or procedural understanding.

The item stem included text and/or graphics. A concerted effort was made to ensure that the language was grade-level appropriate; however, readability statistics were not calculated for each item. Whenever possible, plain language and simple, straightforward statements were incorporated into the items. Graphics were used in instances where they explained the problem, provided a visual clue to clarify the context, or were integral to the stem or answer choices. Irrelevant graphics were not included.

The assessment items were written according to the principles of universal design for assessment (see Ketterlin-Geller, 2005, 2008) and are amenable to accommodations. As delivered, the formative assessment system will include a read aloud feature to support item readability. This ensures that mathematics knowledge and skills are tested, rather than students' reading skills.

The computerized-adaptive test can be administered individually or in a group in an untimed setting.

### **Item Writing Procedures**

---

Item writers were selected based on experience and expertise teaching mathematics at the grade level for which they were selected to write. There were between 5 and 11 item writers per grade. All item writers have between 7 and 19 years of teaching experience. Item writer biographies can be found in Appendix H.

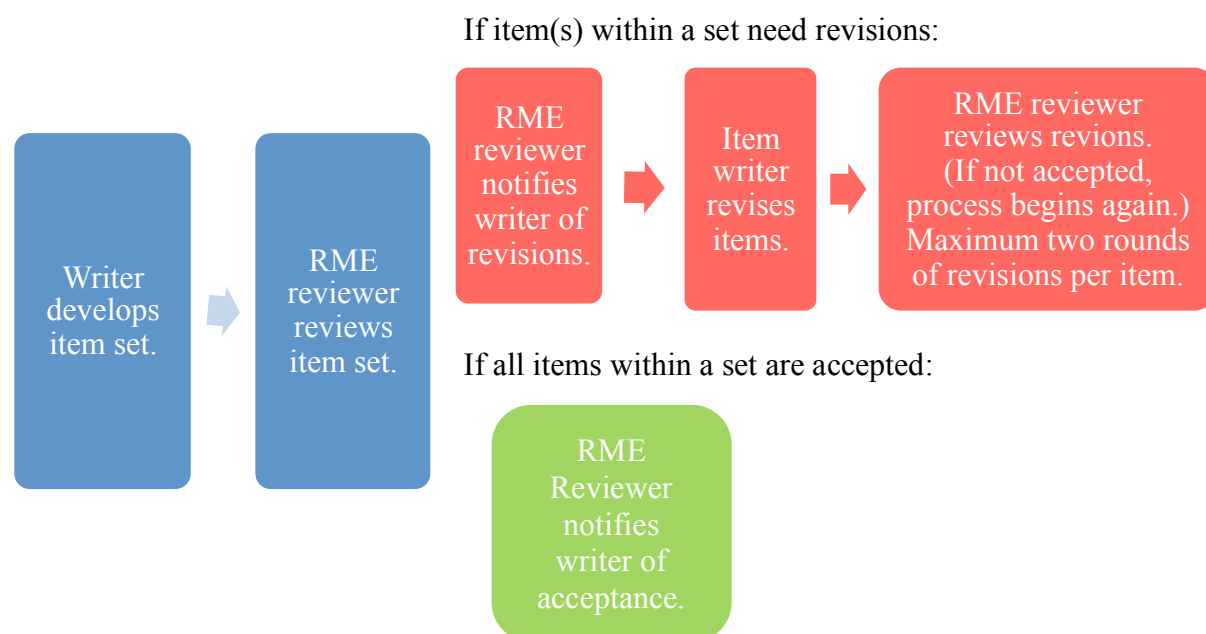
Before beginning the item writing process, item writers engaged in multiple training opportunities. First, all item writers received the RME Style Guide and the Istation Universal Screener Update Item Writer's Guide to review prior to attending virtual training sessions. The style guide provided explanations, examples, and non-examples of stylistic expectations of items to support item writers in writing high-quality mathematics items. The item writer's guide provided an overview of the project, detailed information about the procedures for submission of items and the review process. It also included information on the cognitive levels of engagement and described the principles of universal design for assessment. Second, they participated in two one-hour virtual training sessions. The first training consisted of these topics: (a) overview of the ISIP Mathematics universal screener assessment system and its intended use by teachers; (b) a review of elements of high quality test design as it relates to validity, reliability, and fairness in testing; (c) universal design and universal design for assessment; (d) the levels of cognitive



engagement, specifically the four levels of cognitive engagement used in writing the assessment items: conceptual understanding, procedural fluency, strategic competence, and adaptive reasoning; (e) an overview of the Istation Universal Screener Blueprint.

The second training consisted of these topics: (a) a review of the style guide; and (b) the item writing procedures, which included directions on how to use the data storage system, using the item-writing template, and designing and using graphics in items. At the end of the training, item writers were given an opportunity to ask questions about the process and the expectations outlined.

In preparation for the item writing process, RME staff members assigned item writers a specified number of items to write. Then, each item writer was paired with a staff reviewer. The item writing process was an organic but structured process that promoted collaboration between writers and reviewers. *Figure 1* outlines the item writing process.



*Figure 1*  
*Item writing and review process for all items in the formative assessment item bank*

After items were accepted, RME staff members entered items into an item database. Then, the Istation graphics design team created representational (vector) graphics from the draft (bitmapped) graphics that were made by the item writer. The items were then reviewed externally to evaluate for content-related evidence for validity (described next) and additional recommendations for revisions were made. The final items were copy-edited and reviewed by RME staff and Istation staff.

# Content-Related Evidence for Validity

Mathematics education experts and mathematics teachers evaluated all Grade 2-8 items for accuracy and appropriateness of the content written for the formative assessment item bank.

## Mathematics Education Expert Review

---

Expert reviewers were selected based on current research and teaching in mathematics or mathematics education at the university level. Mathematics education experts were selected to review individual grade levels corresponding with their research and teaching. The education and relevant experience of the expert reviewers are described in the external reviewer biographies in Appendix J.

The mathematics education experts were each asked to review approximately 100 items and evaluate the (a) mathematical accuracy of the content, (b) precision of mathematical vocabulary, and (c) appropriateness of distractors. The criteria used for item evaluation were as follows:

- *Mathematical accuracy of content*: Is the item mathematically accurate?
- *Precision of mathematical vocabulary*: Is the mathematical vocabulary used accurately? Is the mathematical vocabulary precise?
- *Appropriateness of the distractors*: Most students use a process of elimination to narrow their options in the context of multiple-choice questions. The purpose of selecting appropriate distractors is to reduce the likelihood of students with misconceptions from choosing a correct answer in the elimination process. Are the distractors appropriate for the item? Are the distractors mathematically plausible misconceptions?

Items and distractors were evaluated on a 4-point scale for each criterion. A rating of 1 indicated that the item was *Not at all Accurate, Precise, or Appropriate*; a rating of 2 indicated that the item was *Somewhat Accurate, Precise, or Appropriate*; a rating of 3 indicated that the item was *Mostly Accurate, Precise, or Appropriate*; and a rating of 4 indicated the item was *Extremely Accurate, Precise, or Appropriate*. In instances where the reviewer assigned a score of 1 or 2 for any criterion, recommendations were solicited from the expert reviewers that would aid in revisions to an item.

### Grade 2

Overall, the mathematics education experts rated 90% or more of the items as *Extremely Accurate/Appropriate* or *Mostly Accurate/Appropriate* in regards to mathematical content (92%), mathematical vocabulary (93%), and effectiveness of distractors (90%). Their ratings can be seen in Table 2 of Appendix A. For each criteria reviewed, the following paragraph describes how the comments related to items rated with *Somewhat Accurate/Appropriate* and *Not at all Accurate/Appropriate* were addressed.

The mathematics education experts made recommendations for revisions on 47 items. For 11 of the items, the recommendations identified more plausible distractors based on student

misconceptions or the possibility of multiple correct responses. The reconciliation team accepted all of these recommendations and made the changes to the items. For 13 of the items, modifying the language of the question was recommended. These recommendations were based on mathematical precision and the format of the responses. At times, the item referred to a student answering an open-ended prompt, but the responses were numerical. The reconciliation team revised all of the items related to these recommendations. Additional comments from reviewers attended to precision in number lines, place value representations, and other graphics. Revisions to these graphics were submitted to the Istation graphics team. The remaining comments on items referred to correct answers that had been incorrectly coded and mathematically precise vocabulary that needed to be adjusted. These corrections were made to the items.

### *Grade 3*

Overall, the mathematics education experts rated 91% or more of the items as *Extremely Accurate/Appropriate* or *Mostly Accurate/Appropriate* in regards to mathematical content (95%), mathematical vocabulary (96%), and effectiveness of distractors (91%). Their ratings can be seen in Table 2 of Appendix B. For each criteria reviewed, the following paragraphs describe how the feedback was addressed.

The mathematics education experts made recommendations for revisions on 78 items. Of the 78 items, 58 of them received *Extremely Accurate/Appropriate* or *Mostly Accurate/Appropriate* ratings, but the reviewer still gave suggestions for revisions. For 19 of the items with suggestions, the recommendations identified more plausible distractors based on student misconceptions. The reconciliation team accepted 16 of the 19 recommendations and made the recommended changes to the items.

Of the 78 items for which recommendations were made, the mathematics education experts recommended modifications or revisions to the language for 36 of them. These recommendations were based on mathematical precision, tense of the verbs in the item, and spelling errors. The reconciliation team revised 27 of the 36 items related to these recommendations. There were three reasons revisions to the items were not made based on mathematics education experts' recommendations: (1) the suggested wording changes dramatically increased the length of the item, (2) the suggested wording changes included using the word "not" (e.g., Which response is *not* equivalent to this expression?), and/or (3) the distractors did not clearly represent a specific error. In adherence with RME Item Writing and Style Guidelines (Research in Mathematics Education, 2014), we purposefully try to limit the number of words within items to only those necessary to explain the content appropriately and pose a question in an effort to minimize the cognitive burden placed on the student and to minimize the possibility of construct irrelevant variance (i.e., reading skills implicitly assessed on a mathematics tests). In addition, evidence-based guidelines for item writing (Haladyna, Downing, & Rodriguez, 2004) recommend avoidance of negative phrasing in items in an effort to mitigate the possibility that students may misinterpret negatively worded questions. Finally, because this assessment is not designed to provide teachers with information to pinpoint student misconceptions and/or errors associated with the assessed content we agreed that distractors within which integrated skills were represented were acceptable.

For 31 of the items with recommendations, modifications to the graphics were recommended to make the items mathematically accurate. Revisions to these graphics were submitted to the Istation graphics team. Comments about the remaining two items were made because the questions were not viewable inside the Internet browser window. The reconciliation modified the item so that the item stem, graphic, and question were viewable in the browser window.

#### *Grade 4*

Overall, the mathematics education experts rated 91% or more of the items as *Extremely Accurate/Appropriate* or *Mostly Accurate/Appropriate* in regards to mathematical content (92%), mathematical vocabulary (97%), and effectiveness of distractors (91%). Their ratings can be seen in Table 2 of Appendix C. For each criterion reviewed, the following paragraphs describe how the feedback was addressed.

The mathematics education experts made recommendations for revisions on 86 items. Of the 86 items, 45 of them received *Extremely Accurate/Appropriate* or *Mostly Accurate/Appropriate* ratings, but the reviewer still gave suggestions for revisions. For 16 of the items with suggestions, the recommendations identified more plausible distractors based on student misconceptions. The reconciliation team accepted 14 of the 16 recommendations and made the recommended changes to the items.

Of the 86 items for which recommendations were made, the mathematics education experts recommended we modify or revise the language for 50 of them. These recommendations were based on mathematical precision and rewording the question for greater clarity. The reconciliation team revised 43 of the 50 items based on these recommendations. There were three reasons revisions to the items were not made based on mathematics education experts' recommendations: (1) the suggested wording changes changed the tense of the item to past tense, (2) the suggested wording changes included using the word "if" (e.g., If Max needs 7 marbles, how many should he buy?), and/or (3) the suggested wording changes removed a word that made the sentence a phrase. In adherence with RME Item Writing and Style Guidelines (Research in Mathematics Education, 2014) and with best practices of multiple-choice item development (Haladyna, 2000; Haladyna, Downing, & Rodriguez, 2006), we purposefully use present tense in all items (when possible) to minimize confusion and the possibility of construct irrelevant variance introduced by past tense verbs, which often have different forms than their present tense counterparts and avoid conditional language, such as "if...then" statements that may increase the cognitive burden placed on the student to hold one piece of information in working memory (e.g., Max needs 7 marbles) while considering the second part of the phrase (How many should he buy?).

For 13 of the items with recommendations, modifications to the graphics were recommended to make the items mathematically accurate. The reconciliation team submitted revisions to the Istation graphic team for 9 of the 13 items. The revisions that were not made were related to removing the color labels on images. In adherence with the RME Item Writing and Style Guidelines (Research in Mathematics Education, 2014) and Universal Design for Assessment we aim to support students with visual acuities by identifying the color with letters and a key or the color name on any item in which color is necessary to answer the question.

For one item, the mathematical education expert identified that a correct answer was not one of the answer choices. The reconciliation team corrected this error. Comments for two items were made because the question was not viewable inside the Internet browser window. The reconciliation team modified the item so that the item stem, graphic, and question were viewable in the browser window. Additionally, for four items the reviewer noted that the item was not in sequence or was missing from the system. The reconciliation team relied on reviews by mathematics teachers for these items.

### *Grade 5*

Overall, the mathematics education experts rated 89% or more of the items as *Extremely Accurate/Appropriate* or *Mostly Accurate/Appropriate* in regards to mathematical content (89%), mathematical vocabulary (95%), and effectiveness of distractors (93%). Their ratings can be seen in Table 2 of Appendix D. For each criteria reviewed, the following paragraphs describe how the feedback was addressed.

The mathematics education experts made recommendations for revisions on 79 items. Of the 79 items, 18 of them received *Extremely Accurate/Appropriate* or *Mostly Accurate/Appropriate* ratings, however the reviewer also gave suggestions for revisions. Additionally, five of the 79 items received feedback that related to two different categories: effectiveness of distractors and mathematical content. For 15 of the items with suggestions, the recommendations identified more plausible distractors based on common student misconceptions about the content. The reconciliation team accepted 13 of the 15 recommendations and made the recommended changes to the items. The two recommendations that were not accepted did not provide evidence of which distractor should be replaced and/or why the recommendation was a more effective distractor.

Of the 79 items for which recommendations were made, the mathematics education experts recommended modifications or revisions to the language for 43 of them. These recommendations were based on mathematical precision of language, appropriateness of wording for students (e.g., using “trim” or “cut”), and spelling errors. The reconciliation team revised 38 of the 43 items related to these recommendations. Revisions to the five items were not made because the suggested wording changes dramatically increased the length of the item. In adherence with RME Item Writing and Style Guidelines (Research in Mathematics Education, 2014) and research-based guidelines for writing high-quality multiple-choice items (Haladyna, 2000), we purposefully try to limit the number of words within items to only those necessary to explain the content appropriately and pose a question in an effort to minimize (a) the cognitive burden placed on the student and (b) the possibility of construct irrelevant variance (i.e., reading skills implicitly assessed on a mathematics tests).

For ten of the items with recommendations, modifications to the graphics were recommended to make the items mathematically accurate. Revisions to these graphics were submitted to the Istation graphics team. For one of the items, the recommendation was in relation to defining the variables in a formula. The recommended change was made to the item.

For six items, the mathematical education experts identified that there was no correct answer or multiple correct answers. The reconciliation team corrected these errors. Comments for two items were made because the question was not viewable inside the Internet browser window. The

reconciliation team modified the item so that the item stem, graphic, and question were viewable in the browser window. Additionally, for five items the reviewer noted that the item responses could not be seen. This reviewer did not fully rate these five items because the responses were not visible. When the reconciliation team reconciled the comment the item responses were in LaTeX, a typesetting system designed to produce high-quality mathematical equations in proper form, so they noted that it could have been an Internet browser issue. The reconciliation team relied on the mathematics teacher reviewer feedback for these items.

For the remaining two items that had recommended revisions, the reviewer made note of mathematical inaccuracies in the items. These items were rewritten by a member of the reconciliation team and reviewed by two other members of the team using the mathematics education expert and mathematics teacher criteria.

### *Grade 6*

Overall, the mathematics education experts rated 85% or more of the items as *Extremely Accurate/Appropriate* or *Mostly Accurate/Appropriate* in regards to mathematical content (87%), mathematical vocabulary (85%), and effectiveness of distractors (91%). Their ratings can be seen in Table 2 of Appendix E. For each criteria reviewed, the following paragraphs describe how the feedback was addressed.

The mathematics education experts made recommendations for revisions on 97 items. Of the 97 items, 37 of them received *Extremely Accurate/Appropriate* or *Mostly Accurate/Appropriate* ratings, however the reviewer also gave suggestions for revisions. For 11 of the items with suggestions, the recommendations identified more plausible distractors based on student misconceptions. The reconciliation team accepted nine of the 11 recommendations and made the recommended changes to the items. The two remaining recommendations were not accepted because the reconciliation team disagreed with the reviewers' recommendation, as they would make the answer choice in question stand out from the others.

Of the 97 items for which recommendations were made, the mathematics education experts recommended modifications or revisions to the language for 63 of them. These recommendations were based on mathematical precision of language, a more concise way to ask the question, or spelling errors. The reconciliation team revised all but one of the items related to these recommendations. The revision that was not made was the use of the phrase *statistical question* in an item stem. This revision was not made because the language was in the content standard related to the item.

For nine of the items that received recommendations to revise the graphics, revisions were recommended to make the items more mathematically accurate. Revisions to these graphics were submitted to the Istation graphics team. For four of the items, the recommendation was in relation to defining the variables within the stem or changing the variable so that the  $x$  was not in conflict with the multiplication symbol ( $\times$ ). The recommended change was made to the four items. On one item, the reviewer recommended that the formula for surface area be removed because the problem was conceptual. The reconciliation team agreed with this change and removed the formula from the item.

For 10 items, the mathematical education experts identified that there was no correct answer or there were multiple correct answers. The reconciliation team corrected these errors. Comments for two items were made because the question was not viewable inside the Internet browser window. The reconciliation modified the item so that the item stem, graphic, and question were viewable in the browser window.

For one of the financial literacy items the reviewer did not rate the item because he/she did not feel qualified to give a valid rating for the content being assessed. The reconciliation team relied on the mathematics teacher reviewers' feedback for this item.

For the remaining seven items that had recommended revisions, the reviewer made note of mathematical inaccuracies or contexts that were not plausible. These items were rewritten by a member of the reconciliation team and reviewed by two other members of the team using the mathematics education expert and mathematics teacher criteria.

### *Grade 7*

Overall, the mathematics education experts rated 91% or more of the items as *Extremely Accurate/Appropriate* or *Mostly Accurate/Appropriate* in regards to mathematical content (94%), mathematical vocabulary (91%), and effectiveness of distractors (93%). Their ratings can be seen in Table 2 of Appendix F. For each criterion reviewed, the following paragraphs describe how the feedback was addressed.

The mathematics education experts made recommendations for revisions on 62 items. Of the 62 items, 14 of them received *Extremely Accurate/Appropriate* or *Mostly Accurate/Appropriate* ratings, however the reviewer also gave suggestions for revisions. For 13 of the items with suggestions, the recommendations identified more plausible distractors based on student misconceptions. The reconciliation team accepted nine of the 13 recommendations and made the recommended changes to the items. Two recommendations that were not accepted because the reconciliation team disagreed with the reviewers' recommendation, as modification to the response would make it stand out from the other answer choices or did not directly contribute to a student misconception or error (e.g., none of these).

Of the 62 items for which recommendations were made, the mathematics education experts recommended modifications or revisions to the language for 24 of them. These recommendations were based on mathematical precision of language (e.g., identifying that the item stem referred to a constant rate, constant speed, or same cost), a more concise way to ask the question, and spelling errors. The reconciliation team revised all but one of the items related to these recommendations. The revision that was not made was the use of the phrase *without replacement* in a probability item stem. This revision was not made because the language was a part of the stem but italicized for emphasis. The reconciliation team suspected that the Latex did not render when the reviewer was viewing the item.

For four of the items with recommendations, modifications to the graphics were recommended to make the items mathematically accurate. Revisions for three of these items were submitted to the Istation graphics team. For one of the items, the recommendation was to include the value of pi

( $\pi$ ) as a part of the stem language. The reconciliation team added the approximate value of pi ( $\pi \approx 3.14$ ) below the question, but did not include it in the stem text.

For 12 items, the mathematical education experts identified that there was no correct answer or that there were multiple correct answers. The reconciliation team corrected these errors for all but one of the items. The reconciliation team verified with two team members that one of the items did indeed have the correct answer. Comments for two items were made because the question was not viewable inside the window or the inequalities were not visible. For one of these items the reconciliation team modified the item so that the item stem, graphic, and question were viewable in the window. For the other item, the reconciliation team verified that the inequality was visible and suspect that the Latex did not render at the time the reviewer looked at the item.

For the remaining six items that had recommended revisions, the reviewer made note of mathematical inaccuracies or contexts that were not plausible. These items were rewritten by a member of the reconciliation team and reviewed by two other members of the team using the mathematics education expert and mathematics teacher criteria.

### *Grade 8*

Overall, the mathematics education experts rated 91% or more of the items as *Extremely Accurate/Appropriate* or *Mostly Accurate/Appropriate* in regards to mathematical content (91%), mathematical vocabulary (94%), and effectiveness of distractors (92%). Their ratings can be seen in Table 2 of Appendix G. For each criteria reviewed, the following paragraphs describe how the feedback was addressed.

The mathematics education experts made recommendations for revisions to 58 items. Of the 58 items, 13 of them received *Extremely Accurate/Appropriate* or *Mostly Accurate/Appropriate* ratings, however the reviewer also gave suggestions for revisions. For three of the items with suggestions, the recommendations identified more plausible distractors based on student misconceptions. The reconciliation team accepted all three of the recommendations and made the recommended changes to the items.

Of the 58 items for which recommendations were made, the mathematics education experts recommended modifications or revisions to the language of 19 of the items. These recommendations were based on mathematical precision of language, a more concise way to ask the question, or spelling errors. The reconciliation team incorporated all of the recommendations into revisions of the items.

For 12 of the items with recommendations, modifications to the graphics were recommended to make the items more mathematically accurate. Revisions for all but one of these items were submitted to the Istation graphics team. The revision not accepted included adding an equals symbol (=) between the two models in the stem. In adherence with RME Item Writing and Style Guidelines (Research in Mathematics Education, 2014), we avoid using symbols between mathematical models because the models themselves are not equivalent but rather the values represented by the models are equivalent.



Nine comments were related to expressions, equations, and mathematical symbols. The reviewer noted an error in the expression or equation for six of these items. These errors were a consequence of the Latex code used to render the item correctly in the testing platform; these errors were corrected. For one of the items, the recommendation was to change the value of pi ( $\pi$ ) from  $\pi \approx 3$  to  $\pi \approx 3.14$ . Based on the seventh grade content standards, this change was made. For one item, the reviewer recommended removing the formula. This recommendation was accepted on the basis that the item was testing the conceptual understanding of finding the volume of the figure, thus the formula was not necessary. Lastly, one item had two formulas listed. The reviewer recommended labeling the formulas so that the student knew which formula was to be used for each figure. This change was made to increase clarity for students.

For 11 items, the mathematical education experts identified that there was no correct answer or that there were multiple correct answers. The reconciliation team corrected these errors for all of these items. Comments for one item were made because the question was not viewable inside the Internet browser window or the inequalities were not visible. For this, the reconciliation team modified the item so that the item stem, graphic, and question were viewable in the browser window.

For the remaining three items that had recommended revisions, the reviewer made note of mathematical inaccuracies or contexts that were not plausible. These items were rewritten by a member of the reconciliation team and reviewed by two other members of the team using the mathematics education expert and mathematics teacher educator criteria.

## Mathematics Teacher Review

---

Mathematics teachers were selected as reviewers based on experience and expertise teaching mathematics at the grade level for which they were invited to review. All teacher reviewers had between 3 and 21 years of teaching experience. Teacher reviewers' biographies can be found in Appendix I.

Each mathematics teacher was asked to review approximately 100 items and to evaluate the appropriateness of (a) language, (b) mathematical vocabulary, (c) content or concepts, (d) visual representations, and (e) distractors. The criteria used for item evaluation were as follows:

- *Appropriateness of language*: Is the language used in the item appropriate for students in your grade level? Are the question and response options written so that students in your grade level can understand the meaning of the problem?
- *Appropriateness of mathematical vocabulary*: Is the mathematical vocabulary representative of pre-requisite or instructional expectations in your grade level?
- *Appropriateness of content or concepts*: Is the task representative of prerequisite or instructional expectations in your grade level?
- *Appropriateness of visual representation*: Is the visual representation (i.e., graphic, table, image) used in the item appropriate for students in your grade level? Can students in your

grade level understand the meaning of the visual representation? Is the visual representation of the item clear?

- *Appropriateness of the distractors*: Some students use an eliminating process to narrow their options in the context of multiple-choice questions. The purpose of selecting appropriate distractors is to reduce the likelihood of students with misconceptions from choosing a correct answer in the elimination process. Are the distractors appropriate for the item? Do the distractors discriminate between students with specific misconceptions?

The items and distractors were rated on a scale of 1 to 4 for each criterion. A rating of 1 indicated that the item/distractors were *not at all appropriate* based on the instances where the teachers provided a rating of 2 or lower, they were asked to provide additional suggestions and comments to improve the item.

Mathematics teachers also analyzed each item for potential bias in language and/or content. The criterion for potential bias is:

*Bias in language or content*: Does the item require background knowledge unrelated to the concept being tested that would differ for students with different backgrounds? Is the language sensitive to students from diverse backgrounds, students with limited English proficiency and students with special needs? Example: “*What is the most appropriate measurement unit for the length of a sub or hoagie?*” may be unfair for students in certain geographic regions and students with diverse background who are unfamiliar with these terms.

Mathematics teachers were asked to rate each item as *not biased*, *somewhat biased*, or *biased*. In instances where the teachers identified items as *biased*, they were asked to provide additional suggestions and comments to improve the item.

## Grade 2

Overall, the mathematics teachers rated 89% of the items or more as *extremely appropriate* or *mostly appropriate* in regards to mathematical language (89%), mathematical vocabulary (85%), mathematical content (93%), visual representations (91%), and effectiveness of distractors (93%). Their ratings can be seen in Table 3 of Appendix A. For each criteria reviewed, the following paragraph describes how the comments related to items rated with *somewhat appropriate* and *not at all appropriate* were addressed.

For mathematical language, teachers recommended synonyms that would be more accessible for second graders (e.g., *shows* instead of *represent*). These changes were accepted in instances where using a synonym did not impact the mathematical vocabulary assessed by the content standard. For mathematical vocabulary, most of the comments referred to the teacher’s recommendations that *number sentence* be used in any problem that used the mathematical vocabulary of *expression* or *equation*. The reconciliation team reviewed the mathematical content standards and determined that it was appropriate to use *expression*, *equation*, and *number sentence* in second grade. Each item with this mathematical vocabulary was reviewed and verified that the appropriate word was used (e.g., *expression* vs. *equation*). Other mathematical

vocabulary recommendations were reviewed in the same manner by analyzing the appropriateness based on the mathematical content standards.

For mathematical content, the teachers commented on seven items that contained inappropriate content for second grade according to the Texas Essential Knowledge and Skills (TEKS). For those items, the integrated standards document was referenced and the items were aligned to the CCSS-M, rather than the TEKS. These items were not modified. For visual representations, the teachers identified 15 items in which graphics could be enhanced. Revisions to these graphics were submitted to the Istation graphics team. Finally, the reviewers noted 11 items in which the identified correct answer was incorrect or the distractors were not appropriate. Five of these items were significantly revised based on the teacher reviewers' recommendations. For the remaining six, the correct answer had been incorrectly coded and was corrected.

Mathematics teachers identified 2 items as "somewhat biased." Reviewers made recommendations of how the items could be revised to remove the bias. The recommended changes were made to the items. Table 4 of Appendix A shows each reviewer's ratings.

### *Grade 3*

Overall, the mathematics teachers rated 97% or more of the items as *Extremely Appropriate* or *Mostly Appropriate* in regards to mathematical language (97%), mathematical vocabulary (99%), mathematical content (99%), visual representations (99%), and effectiveness of distractors (99%). Their ratings can be seen in Table 3 of Appendix B. For each criteria reviewed, the following paragraph describes how the comments related to items rated with *Somewhat Appropriate* and *Not at all Appropriate* were addressed.

For mathematical language and mathematical vocabulary, teachers recommended that the wording of the items be revised to increase readability for third graders. These changes were accepted in instances where the recommendation did not impact the mathematical vocabulary assessed by the content standard.

For mathematical content, the teachers commented that three items contained inappropriate content for third grade according to the TEKS. For those items, the integrated standards document was referenced and the items were aligned to the CCSS-M, rather than the TEKS. These items were not modified. Additionally, an item originally written to the *adaptive* cognitive engagement level was modified to be *conceptual* to make the mathematical content more appropriate for the grade level.

For visual representations, the teachers identified nine items in which graphics could be enhanced. Revisions to graphics for six of the nine items were submitted to the Istation graphics team; the remaining three items were not modified due to an alternate recommendation from the mathematics education expert. For effectiveness of distractors, three suggestions were given, of which two were accepted. Finally, two items with comments indicated that the question was not viewable inside the Internet browser window. The reconciliation modified the item so that the item prompt, graphic, and question were viewable in the window.

Mathematics teachers rated one item as *Somewhat Biased*. The reviewer made recommendations as to how the item could be revised to eliminate that bias. The recommended change was made to the item. Table 4 of Appendix B shows each reviewer's ratings.

#### *Grade 4*

Overall, the mathematics teachers rated 83% or more of the items as *Extremely Appropriate* or *Mostly Appropriate* in regards to mathematical language (97%), mathematical vocabulary (91%), mathematical content (83%), visual representations (94%), and effectiveness of distractors (96%). Their ratings can be seen in Table 3 of Appendix C. For each criteria reviewed, the following paragraph describes how the comments related to items rated with *Somewhat Appropriate* and *Not at all Appropriate* were addressed.

For mathematical language and mathematical vocabulary, teachers recommended that the wording of the questions within the stem be revised to support students in clearly identifying what the question was asking. These changes were accepted in instances where the recommendation did not impact the mathematical vocabulary assessed by the content standard.

For mathematical content, 44 of the 47 items that were rated as *Somewhat Appropriate* and *Not at all Appropriate* were identified by mathematics teachers as containing inappropriate content for third grade according to the TEKS. For those items, the integrated standards document was referenced and the items were aligned to the CCSS-M, rather than the TEKS. These items were not modified. The remaining three items with recommendations related to mathematical content were accepted and the item was modified to address the concern.

For visual representations, the teachers identified seven items in which graphics could be enhanced. Revisions to graphics for six of the seven recommendations were submitted to the Istation graphic team; the remaining item was not modified because the right angle symbol was necessary for the student to accurately solve the problem without visually having to discern that the angle was ninety degrees. For effectiveness of distractors, three suggestions were given, of which all three were accepted.

On two items comments indicated that the question was not viewable inside the Internet browser window. The reconciliation modified the item so that the item prompt, graphic, and question were viewable in the browser window. For three items, the reviewer identified that a correct answer was not one of the answer choices or multiple correct answers were plausible. The reconciliation team corrected these errors. Finally, for 22 items the reviewer noted that the item was not in sequence (i.e., the items did not appear in the order the reviewer anticipated based on the Excel spreadsheet) or missing from the system. The reconciliation team relied on the mathematics education expert feedback for these items.

Mathematics teachers rated two items as *Somewhat Biased* and one item as *biased*. The reviewer made recommendations as to how the item could be revised to eliminate that bias. The recommended changes were made to the item. Table 4 of Appendix C shows each reviewer's ratings.

## Grade 5

Overall, the mathematics teachers rated 87% or more of the items as *Extremely Appropriate* or *Mostly Appropriate* in regards to mathematical language (97%), mathematical vocabulary (92%), mathematical content (87%), visual representations (91%), and effectiveness of distractors (95%). Their ratings can be seen in Table 3 of Appendix D. For each criteria reviewed, the following paragraphs describe how the comments related to items rated with *Somewhat Appropriate* and *Not at all Appropriate* were addressed.

For mathematical language, teachers recommended that the wording of 15 items be modified so that the construct was clearer for students. These changes were accepted in instances where the recommendation did not impact the mathematical vocabulary assessed by the content standard. Additionally, teachers identified two items with spelling mistakes. These mistakes were corrected.

For mathematical vocabulary, teachers recommended that four items have words replaced with synonyms (e.g., cook instead of chef) or more mathematically appropriate vocabulary (e.g., expression instead of equation). Four of the five recommended revisions were made. The one recommended revision that was not made was to replace the word *strategy* with *expression*; the reconciliation team agreed that *strategy* was appropriate language for the item.

For mathematical content, 33 of the 40 items that were rated as *Somewhat Appropriate* and *Not at all Appropriate* were identified as containing inappropriate content for fifth grade according to the TEKS. For those items, the integrated standards document was referenced, indicating that the items were aligned to the CCSS-M, rather than the TEKS. These items were not modified. Recommendations for the remaining six items related to mathematical content were accepted and these items were modified to address the reviewers' concerns.

For visual representations, the teachers identified 28 items in which graphics could be enhanced or were not viewable in their Internet browser. Revisions to graphics were recommended for 13 of the 28 items. These revisions were submitted to the Istation graphic team. The remaining 15 items did not require any modifications because the reviewer commented that the low rating was because the item did not have a visual representation. The reviewer did not make any recommendations for adding visual representations to the item so the reconciliation team interpreted that the reviewer thought any item without a visual representation should receive a low rating. For effectiveness of distractors, 10 suggestions were given, of which nine were made. For the item in which the revisions to the distractors were not made, the reconciliation team modified the language of the stem to address the teacher's concern.

On two items, comments indicated that the question was not viewable inside the Internet browser window. The reconciliation team modified the item so that the item prompt, graphic, and question were viewable in the window. For eight items, a reviewer identified that there was not a correct answer or that there were multiple correct answers. The reconciliation team corrected these errors.

Mathematics teacher reviewers rated one item as *Somewhat Biased* and one item as *Biased*. The reviewers made recommendations as to how the item could be revised to eliminate that bias. The

recommended changes were made to the item. Table 4 of Appendix D shows each reviewer's ratings.

### *Grade 6*

Overall, the mathematics teacher reviewers rated 93% or more of the items as *Extremely Appropriate* or *Mostly Appropriate* in regards to mathematical language (97%), mathematical vocabulary (98%), mathematical content (96%), visual representations (97%), and effectiveness of distractors (93%). Their ratings can be seen in Table 3 of Appendix E. For each criteria reviewed, the following paragraph describes how the comments related to items rated with *Somewhat Appropriate* and *Not at all Appropriate* were addressed.

For mathematical language, teachers recommended that the wording of eight items be modified so that the construct was clearer for students. These changes were accepted in instances where the recommendation did not impact the mathematical vocabulary assessed by the content standard. Additionally, teachers identified three items with spelling mistakes. These mistakes were corrected.

For mathematical vocabulary, teachers recommended that five items be revised to include more grade-level appropriate or more mathematically appropriate vocabulary (e.g., *expression* instead of *equation*). All of the recommended changes were made. Additionally, a teacher commented that one item was "wordy." The reconciliation team reviewed the content standard and level of cognitive engagement before determining that the item could not be further modified without changing the intended construct of the item. Finally, for mathematical vocabulary a teacher reviewer identified two instances where a variable should be adjusted. In one instance the multiplication sign (x) and the variable  $x$ , were both used in an item. In the other instance the variable  $s$  was used for sweater but the item also referred to shirts. For both of these items a modification was made to eliminate confusion for students.

For mathematical content, six of the 12 items that mathematics teachers rated as *Somewhat Appropriate* and *Not at all Appropriate* were identified as containing inappropriate content or formulas for sixth grade as detailed in the TEKS. For those items, the integrated standards document was referenced and the items were aligned to the CCSS-M, rather than the TEKS. These items were not modified. The remaining six items with recommendations related to mathematical content were accepted and the item was modified to address the concern.

For visual representations, the teachers identified nine items in which graphics should be modified/enhanced. Additionally, adding a visual representation to one item was recommended. Revisions to graphics were recommended for all 10 items. These revisions were submitted to the Istation graphics team.

On one item, comments indicated that the question was not viewable inside the Internet browser window. The reconciliation modified the item so that the item prompt, graphic, and question were viewable in the browser window. For 16 items, a reviewer identified that there was not a correct answer or that there were multiple correct answers. The reconciliation team corrected these errors. Additionally, for five items reviewers recommended a distractor be replaced with a more plausible distractor that was based on common student misconceptions about the content.

Rationales for each of the five recommendations were given and the reconciliation team adjusted the items.

Mathematics teacher reviewers rated six items as *Somewhat Biased* and one item as *Biased*. The reviewers made recommendations as to how the item could be revised to eliminate that bias. The recommended changes were made to the items. Table 4 of Appendix E shows each reviewer's ratings.

### *Grade 7*

Overall, the mathematics teachers rated 91% or more of the items as *Extremely Appropriate* or *Mostly Appropriate* in regard to mathematical language (96%), mathematical vocabulary (96%), mathematical content (91%), visual representations (95%), and effectiveness of distractors (99%). Their ratings can be seen in Table 3 of Appendix F. For each criteria reviewed, the following paragraph describes how the comments related to items rated with *Somewhat Appropriate* and *Not at all Appropriate* were addressed.

For mathematical language, teachers recommended that the wording of 12 items be modified so that the construct was clearer for students. These changes were accepted in instances where the recommendation did not impact the mathematical vocabulary assessed by the content standard. Additionally, teachers identified two items with spelling mistakes. These mistakes were corrected.

For mathematical vocabulary, teachers rated 11 items as *Somewhat Appropriate* and *Not at all Appropriate*. The ratings for this category were in conjunction with low ratings for other categories. The reviewers did not specifically comment on mathematical vocabulary, thus the concerns were addressed in other ways.

For mathematical content, 28 items were rated as *Somewhat Appropriate* and *Not at all Appropriate*. Of these 28 items only five of the comments specifically related to mathematical content. Modifications to these items were made to address the reviewers concerns. An additional seven items rated *Somewhat Appropriate* and *Not at all Appropriate* were identified by mathematics teachers as containing inappropriate content for seventh grade as detailed in the TEKS. For those items, the integrated standards document was referenced and the items were aligned to the CCSS-M, rather than the TEKS. These items were not modified. For the remaining 16 items with recommendations not related to mathematical content, we addressed reviewers' concerns by attending to other features of the items.

For visual representations, the reviewers identified seven items in which graphics should be modified or enhanced. These revisions were submitted to the Istation graphic team. The remaining eight items rated as *Somewhat Appropriate* and *Not at all Appropriate* for visual representations did not have a comment directly related to the visual representation, thus, the concerns were addressed in other ways. For visual representations, the reviewers rated three items as *Not at all Appropriate* but did not give any rationale for the rating so no modification was made to the items.

Mathematics teachers rated seven items as *Somewhat Biased* and three items as *Biased*. The reviewers made recommendations on how to eliminate the bias context on five of these 10 items.

The recommended changes were made to these items. For the remaining 5 items recommendations to eliminate bias were not provided; the reconciliation team consulted and determined that revisions to those items were not necessary. Table 4 of Appendix F shows each reviewer's ratings.

### *Grade 8*

Overall, the mathematics teachers rated 95% or more of the items as *Extremely Appropriate* or *Mostly Appropriate* in regard to mathematical language (96%), mathematical vocabulary (97%), mathematical content (95%), visual representations (95%), and effectiveness of distractors (97%). Their ratings can be seen in Table 3 of Appendix G. For each criteria reviewed, the following paragraph describes how the comments related to items rated as *Somewhat Appropriate* and *Not at all Appropriate* were addressed.

For mathematical language, teachers recommended that the wording of 11 items be modified so that the construct was clearer for students. These changes were accepted in instances where the recommendation did not impact the mathematical vocabulary assessed by the content standard. For two items, the recommendations impacted the content assessed by the standard and the revisions were not made.

For mathematical vocabulary, teachers recommended that eight items be revised to include more grade-level appropriate, or more mathematically appropriate vocabulary. All of the recommended changes were made as part of the changes for mathematical language comments.

For mathematical content, one of the 14 items that reviewers rated as *Somewhat Appropriate* was identified as containing inappropriate content or formulas for eighth grade as outlined by the TEKS. The integrated standards document was referenced and the items were aligned to the CCSS-M, rather than the TEKS. This item was not modified. Of the remaining 13 items with recommendations related to mathematical content, eight of the recommended changes were addressed by attending to comments made about the mathematical language of the items. Three of the remaining comments were not addressed because the items were designed to elicit students' conceptual understanding of the content and the recommended revisions would have restructured the item to elicit students' procedural fluency. Lastly, two comments were not necessary changes because they pertained to not being able to visually see the expressions in the responses. The reconciliation team suspects that the reviewer was not able to see the content because of an Internet issue at the time these two items were viewed. Both of the items had the expressions as a part of the item when the reconciliation team reviewed the items.

For visual representations, the teachers identified 16 items in which graphics should be modified or enhanced. Revisions were submitted to the Istation graphics team for 13 of the items. For the remaining three ratings of *Somewhat Appropriate* and *Not at all Appropriate*, a recommendation directly related to the visual representation was not provided. The reconciliation team believes the concern was addressed through the modifications made in the other categories of the review.

For 11 items, a reviewer identified that there was not a correct answer or that there were multiple correct answers. The reconciliation team corrected these errors. Additionally, for eight items the reviewer recommended a distractor be replaced with another more plausible distractor based on



common student misconceptions about the content. Rationale for each of the eight recommendations was given and the reconciliation team adjusted the items.

Mathematics teacher reviewers rated three items as *Somewhat Biased* and two items as *Biased*. The reviewers made recommendations as to how the item could be revised to eliminate that bias. The recommended changes were made to the items. Table 4 of Appendix G shows each reviewer's ratings.

## Conclusion

The purpose of this technical report was to describe the development of an updated item bank for the Grade 2 through Grade 8 formative assessment for the ISIP Mathematics universal screeners. We described the components of the assessed construct—content standards and levels of cognitive complexity—and the process for sampling the content assessed in the formative assessment item bank. Next, we described the item writing procedures and provided the qualifications for the item writers. Finally, we documented the process and outcomes of an external item review conducted by mathematics education experts and mathematics teachers to contribute to content-related evidence for validity.

---

## References

---

- American Educational Research Association (AERA), American Psychological Association (APA), & National Council on Measurement in Education (NCME). (1999). *Standards for educational and psychological testing*. Washington, DC: American Psychological Association.
- Haladyna, T. M., Downing, S. M., Rodriguez, M.C. (2004). A review of multiple-choice item-writing guidelines for classroom assessment. *Applied Measurement in Education*, 15, 309-333.
- Ketterlin-Geller, L.R. (2005). Knowing what all students know: Procedures for developing universally designed assessments. *Journal of Technology, Learning, and Assessment*, 4(2). Available from <http://ejournals.bc.edu/ojs/index.php/jtla/issue/view/188>
- Ketterlin-Geller, L.R. (2008). Testing students with special needs: A model for understanding the interaction between assessment and student characteristics in a universally designed environment. *Educational Measurement: Issues and Practice*, 27(3), 3-16. DOI: 10.1111/j. 1745-3992.2008.00124.x
- National Research Council (2001). *Adding it up: Helping children learn mathematics*. J. Kilpatrick, J. Swafford, & B. Findell (Eds.). Mathematics Learning Study Committee, Center for Education, Division of Behavioral and Social Sciences and Education. Washington, DC: National Academy Press.
- Research in Mathematics Education. (2014). *Research in Mathematics Education Item Writing Style Guide*. Dallas, TX: Southern Methodist University.

## Appendix A: Grade 2 External Review Data

*Table 1: Content Sampling Matrix Grade 2*

Cognitive Engagement Level	Easy	Medium	Difficult	Total
Procedural Fluency	28	30	22	70
Conceptual Understanding	48	44	30	122
Strategic Competence	10	23	20	53
Adaptive Reasoning	8	19	18	45
Total	90	94	116	300

*Table 2: Mathematics Education Expert Review Ratings Grade 2*

Mathematics Education Expert ID	Scale	Mathematical Content	Mathematical Vocabulary	Distractors
1*	4 – Extremely Accurate / Appropriate	83	79	68
	3 – Mostly Accurate / Appropriate	7	14	25
	2 – Somewhat Accurate / Appropriate	8	6	5
	1 – Not at All Accurate / Appropriate	1	0	1
2*	4 – Extremely Accurate / Appropriate	85	80	81
	3 – Mostly Accurate / Appropriate	6	12	10
	2 – Somewhat Accurate / Appropriate	6	5	4
	1 – Not at All Accurate / Appropriate	2	2	4
3	4 – Extremely Accurate / Appropriate	90	86	82
	3 – Mostly Accurate / Appropriate	5	9	4
	2 – Somewhat Accurate / Appropriate	3	5	14
	1 – Not at All Accurate / Appropriate	2	0	0
Total	4 – Extremely Accurate / Appropriate	258	245	231
	3 – Mostly Accurate / Appropriate	18	35	39
	2 – Somewhat Accurate / Appropriate	17	16	23
	1 – Not at All Accurate / Appropriate	5	2	5

\*Mathematics Education Expert 1 and 2 did not review 1 of the 100 items in their set.

Table 3: Mathematics Teacher Review Ratings Grade 2

Teacher Reviewer ID	Scale	Mathematical Language	Mathematical Vocabulary	Mathematical Content	Visual Representations	Distractors
1	4 – Extremely Appropriate	78	70	92	96	98
	3 – Mostly Appropriate	0	0	0	0	0
	2 – Somewhat Appropriate	19	30	2	0	0
	1 – Not at all Appropriate	3	0	6	4	2
2	4 – Extremely Appropriate	76	77	75	87	73
	3 – Mostly Appropriate	20	18	18	6	15
	2 – Somewhat Appropriate	2	4	6	4	8
	1 – Not at all Appropriate	2	1	1	3	4
3	4 – Extremely Appropriate	87	81	89	79	84
	3 – Mostly Appropriate	7	9	6	6	8
	2 – Somewhat Appropriate	2	2	1	6	1
	1 – Not at all Appropriate	4	8	4	9	7
Total	4 – Extremely Appropriate	241	228	256	262	255
	3 – Mostly Appropriate	27	27	24	12	23
	2 – Somewhat Appropriate	23	36	9	10	9
	1 – Not at all Appropriate	9	9	11	16	13

Table 4: Mathematics Teacher Review Ratings for Bias Grade 2

Teacher Reviewer ID	No Bias	Some Bias	Bias
1	100	0	0
2	98	2	0
3	100	0	0
Total	298	2	0

## Appendix B: Grade 3 External Review Data

*Table 1: Content Sampling Matrix Grade 3*

<b>Cognitive Engagement Level</b>	<b>Easy</b>	<b>Medium</b>	<b>Difficult</b>	<b>Total</b>
Procedural Fluency	31	54	17	102
Conceptual Understanding	55	26	25	106
Strategic Competence	10	27	22	59
Adaptive Reasoning	4	9	20	33
<b>Total</b>	<b>100</b>	<b>116</b>	<b>84</b>	<b>300</b>

*Table 2: Mathematics Education Expert Review Ratings Grade 3*

<b>Mathematics Education Expert ID</b>	<b>Scale</b>	<b>Mathematical Content</b>	<b>Mathematical Vocabulary</b>	<b>Distractors</b>
1	4 – Extremely Accurate / Appropriate	81	76	69
	3 – Mostly Accurate / Appropriate	12	20	18
	2 – Somewhat Accurate / Appropriate	5	2	11
	1 – Not at All Accurate / Appropriate	2	2	2
2	4 – Extremely Accurate / Appropriate	98	97	95
	3 – Mostly Accurate / Appropriate	2	1	4
	2 – Somewhat Accurate / Appropriate	0	1	1
	1 – Not at All Accurate / Appropriate	0	1	0
3	4 – Extremely Accurate / Appropriate	90	89	83
	3 – Mostly Accurate / Appropriate	2	5	4
	2 – Somewhat Accurate / Appropriate	2	2	5
	1 – Not at All Accurate / Appropriate	6	4	8
<b>Total</b>	4 – Extremely Accurate / Appropriate	269	262	247
	3 – Mostly Accurate / Appropriate	16	26	26
	2 – Somewhat Accurate / Appropriate	7	5	17
	1 – Not at All Accurate / Appropriate	8	7	10

*Table 3: Mathematics Teacher Review Ratings Grade 3*

Teacher Reviewer ID	Scale	Mathematical Language	Mathematical Vocabulary	Mathematical Content	Visual Representations	Distractors
1	4 – Extremely Appropriate	95	99	97	95	97
	3 – Mostly Appropriate	1	0	0	3	2
	2 – Somewhat Appropriate	4	1	3	1	1
	1 – Not at all Appropriate	0	0	0	1	0
2	4 – Extremely Appropriate	87	87	87	87	85
	3 – Mostly Appropriate	12	12	12	12	14
	2 – Somewhat Appropriate	1	1	1	1	1
	1 – Not at all Appropriate	0	0	0	0	0
3	4 – Extremely Appropriate	95	286	100	100	100
	3 – Mostly Appropriate	2	12	0	0	0
	2 – Somewhat Appropriate	2	2	0	0	0
	1 – Not at all Appropriate	1	0	0	0	0
Total	4 – Extremely Appropriate	277	286	284	282	282
	3 – Mostly Appropriate	15	12	12	15	16
	2 – Somewhat Appropriate	7	2	4	2	2
	1 – Not at all Appropriate	1	0	0	1	0

*Table 4: Mathematics Teacher Review Ratings for Bias Grade 3*

Teacher Reviewer ID	No Bias	Some Bias	Bias
1	100	0	0
2	100	0	0
3*	99	1	0
Total	300	1	0

## Appendix C: Grade 4 External Review Data

*Table 1: Content Sampling Matrix Grade 4*

Cognitive Engagement Level	Easy	Medium	Difficult	Total
Procedural Fluency	24	34	33	91
Conceptual Understanding	55	32	23	110
Strategic Competence	7	27	21	55
Adaptive Reasoning	7	14	23	44
Total	93	107	100	300

*Table 2: Mathematics Education Expert Review Ratings Grade 4*

Mathematics Education Expert ID	Scale	Mathematical Content	Mathematical Vocabulary	Distractors
1*	4 – Extremely Accurate / Appropriate	71	76	58
	3 – Mostly Accurate / Appropriate	9	18	20
	2 – Somewhat Accurate / Appropriate	12	3	18
	1 – Not at All Accurate / Appropriate	7	2	3
2**	4 – Extremely Accurate / Appropriate	94	96	92
	3 – Mostly Accurate / Appropriate	1	0	0
	2 – Somewhat Accurate / Appropriate	0	0	0
	1 – Not at All Accurate / Appropriate	1	0	4
3	4 – Extremely Accurate / Appropriate	85	82	87
	3 – Mostly Accurate / Appropriate	10	14	10
	2 – Somewhat Accurate / Appropriate	4	3	2
	1 – Not at All Accurate / Appropriate	1	1	1
Total	4 – Extremely Accurate / Appropriate	250	254	237
	3 – Mostly Accurate / Appropriate	20	32	30
	2 – Somewhat Accurate / Appropriate	16	6	20
	1 – Not at All Accurate / Appropriate	9	3	8

\*Mathematics Education Expert 1 did not review 1 of the 100 items in their set.

\*Mathematics Education Expert 2 did not review 4 of the 100 items in their set.



*Table 3: Mathematics Teacher Review Ratings Grade 4*

Teacher Reviewer ID	Scale	Mathematical Language	Mathematical Vocabulary	Mathematical Content	Visual Representations	Distractors
1	4 – Extremely Appropriate	87	99	96	94	84
	3 – Mostly Appropriate	9	1	2	3	12
	2 – Somewhat Appropriate	2	0	2	2	1
	1 – Not at all Appropriate	2	0	0	1	3
2	4 – Extremely Appropriate	98	76	70	93	96
	3 – Mostly Appropriate	1	1	3	0	1
	2 – Somewhat Appropriate	0	12	12	2	1
	1 – Not at all Appropriate	1	11	15	5	2
3*	4 – Extremely Appropriate	32	41	41	59	38
	3 – Mostly Appropriate	44	34	19	13	36
	2 – Somewhat Appropriate	2	2	8	6	2
	1 – Not at all Appropriate	0	1	10	0	2
Total	4 – Extremely Appropriate	217	216	207	246	218
	3 – Mostly Appropriate	54	36	24	16	49
	2 – Somewhat Appropriate	4	14	22	10	4
	1 – Not at all Appropriate	3	12	25	6	7

\*Mathematics Teacher 3 did not review 22 of the 100 items in their set.

*Table 4: Mathematics Teacher Review Ratings for Bias Grade 4*

Teacher Reviewer ID	No Bias	Some Bias	Bias
1	97	2	1
2	100	0	0
3	78	0	0
Total	275	2	1

\*Mathematics Teacher 3 did not review 22 of the 100 items in their set.

## Appendix D: Grade 5 External Review Data

*Table 1: Content Sampling Matrix Grade 5*

Cognitive Engagement Level	Easy	Medium	Difficult	Total
Procedural Fluency	30	34	25	89
Conceptual Understanding	33	39	26	98
Strategic Competence	25	24	26	75
Adaptive Reasoning	11	16	11	38
Total	99	113	88	300

*Table 2: Mathematics Education Expert Review Ratings Grade 5*

Mathematics Education Expert ID	Scale	Mathematical Content	Mathematical Vocabulary	Distractors
1	4 – Extremely Accurate / Appropriate	97	100	99
	3 – Mostly Accurate / Appropriate	0	0	0
	2 – Somewhat Accurate / Appropriate	0	0	0
	1 – Not at All Accurate / Appropriate	3	0	1
2*	4 – Extremely Accurate / Appropriate	62	65	65
	3 – Mostly Accurate / Appropriate	17	19	17
	2 – Somewhat Accurate / Appropriate	12	11	13
	1 – Not at All Accurate / Appropriate	9	3	0
3**	4 – Extremely Accurate / Appropriate	84	89	85
	3 – Mostly Accurate / Appropriate	3	4	4
	2 – Somewhat Accurate / Appropriate	4	1	3
	1 – Not at All Accurate / Appropriate	4	1	3
Total	4 – Extremely Accurate / Appropriate	243	254	249
	3 – Mostly Accurate / Appropriate	20	23	21
	2 – Somewhat Accurate / Appropriate	16	12	16
	1 – Not at All Accurate / Appropriate	16	4	4

\* Mathematics Education Expert 2 did not rate 5 of the 100 items in their set for distractors and did not rate 2 of the 100 items in their set for mathematical vocabulary

\*\* Mathematics Education Expert 3 did not review 5 of the 100 items in their set.

*Table 3: Mathematics Teacher Review Ratings Grade 5*

Teacher Reviewer ID	Scale	Mathematical Language	Mathematical Vocabulary	Mathematical Content	Visual Representations	Distractors
1	4 – Extremely Appropriate	62	61	49	55	54
	3 – Mostly Appropriate	34	35	32	24	39
	2 – Somewhat Appropriate	4	4	7	20	6
	1 – Not at all Appropriate	0	0	12	1	1
2*	4 – Extremely Appropriate	89	86	89	91	90
	3 – Mostly Appropriate	7	0	0	3	6
	2 – Somewhat Appropriate	1	3	0	4	3
	1 – Not at all Appropriate	3	11	11	2	0
3	4 – Extremely Appropriate	81	90	77	94	69
	3 – Mostly Appropriate	18	5	13	5	25
	2 – Somewhat Appropriate	1	5	1	1	6
	1 – Not at all Appropriate	0	0	9	0	0
Total	4 – Extremely Appropriate	232	237	215	240	213
	3 – Mostly Appropriate	59	40	45	32	70
	2 – Somewhat Appropriate	6	12	8	25	15
	1 – Not at all Appropriate	3	11	32	3	1

\*Mathematics Teacher 3 did not fully review 1 of the 100 items in their set.

*Table 4: Mathematics Teacher Review Ratings for Bias Grade 5*

Teacher Reviewer ID	No Bias	Some Bias	Bias
1	100	0	0
2	99	1	0
3	99	0	1
Total	298	1	1

## Appendix E: Grade 6 External Review Data

*Table 1: Content Sampling Matrix Grade 6*

<b>Cognitive Engagement Level</b>	<b>Easy</b>	<b>Medium</b>	<b>Difficult</b>	<b>Total</b>
Procedural Fluency	26	28	33	87
Conceptual Understanding	47	32	27	106
Strategic Competence	17	26	21	64
Adaptive Reasoning	12	13	18	43
<b>Total</b>	<b>102</b>	<b>99</b>	<b>99</b>	<b>300</b>

*Table 2: Mathematics Education Expert Review Ratings Grade 6*

<b>Mathematics Education Expert ID</b>	<b>Scale</b>	<b>Mathematical Content</b>	<b>Mathematical Vocabulary</b>	<b>Distractors</b>
1	4 – Extremely Accurate / Appropriate	91	88	94
	3 – Mostly Accurate / Appropriate	1	1	1
	2 – Somewhat Accurate / Appropriate	7	10	5
	1 – Not at All Accurate / Appropriate	1	1	0
2	4 – Extremely Accurate / Appropriate	55	54	66
	3 – Mostly Accurate / Appropriate	14	18	14
	2 – Somewhat Accurate / Appropriate	13	24	15
	1 – Not at All Accurate / Appropriate	16	2	3
3	4 – Extremely Accurate / Appropriate	100	93	94
	3 – Mostly Accurate / Appropriate	1	2	5
	2 – Somewhat Accurate / Appropriate	0	6	3
	1 – Not at All Accurate / Appropriate	1	1	0
<b>Total</b>	4 – Extremely Accurate / Appropriate	246	235	249
	3 – Mostly Accurate / Appropriate	16	21	20
	2 – Somewhat Accurate / Appropriate	20	40	23
	1 – Not at All Accurate / Appropriate	18	4	3

*Table 3: Mathematics Teacher Review Ratings Grade 6*

Teacher Reviewer ID	Scale	Mathematical Language	Mathematical Vocabulary	Mathematical Content	Visual Representations	Distractors
1	4 – Extremely Appropriate	96	95	94	93	92
	3 – Mostly Appropriate	2	1	1	3	1
	2 – Somewhat Appropriate	0	2	3	0	4
	1 – Not at all Appropriate	0	0	0	2	1
2	4 – Extremely Appropriate	96	99	91	93	87
	3 – Mostly Appropriate	4	1	2	4	5
	2 – Somewhat Appropriate	0	0	4	0	5
	1 – Not at all Appropriate	0	0	3	3	3
3	4 – Extremely Appropriate	0	0	0	1	1
	3 – Mostly Appropriate	94	98	100	96	93
	2 – Somewhat Appropriate	8	3	2	5	8
	1 – Not at all Appropriate	0	1	0	0	0
Total	4 – Extremely Appropriate	192	194	185	187	180
	3 – Mostly Appropriate	100	100	103	103	99
	2 – Somewhat Appropriate	8	5	9	5	17
	1 – Not at all Appropriate	0	1	3	5	4

*Table 4: Mathematics Teacher Review Ratings for Bias Grade 6*

Teacher Reviewer ID	No Bias	Some Bias	Bias
1	94	3	1
2	97	3	0
3	102	0	0
Total	293	6	1

## Appendix F: Grade 7 External Review Data

*Table 1: Content Sampling Matrix Grade 7*

<b>Cognitive Engagement Level</b>	<b>Easy</b>	<b>Medium</b>	<b>Difficult</b>	<b>Total</b>
Procedural Fluency	35	41	30	106
Conceptual Understanding	37	32	24	93
Strategic Competence	26	23	16	65
Adaptive Reasoning	9	10	17	36
<b>Total</b>	<b>107</b>	<b>106</b>	<b>87</b>	<b>300</b>

*Table 2: Mathematics Education Expert Review Ratings Grade 7*

<b>Mathematics Education Expert ID</b>	<b>Scale</b>	<b>Mathematical Content</b>	<b>Mathematical Vocabulary</b>	<b>Distractors</b>
1	4 – Extremely Accurate / Appropriate	96	88	99
	3 – Mostly Accurate / Appropriate	0	1	0
	2 – Somewhat Accurate / Appropriate	4	11	1
	1 – Not at All Accurate / Appropriate	0	0	0
2	4 – Extremely Accurate / Appropriate	85	85	77
	3 – Mostly Accurate / Appropriate	3	2	4
	2 – Somewhat Accurate / Appropriate	5	10	19
	1 – Not at All Accurate / Appropriate	7	3	0
3	4 – Extremely Accurate / Appropriate	98	97	96
	3 – Mostly Accurate / Appropriate	0	1	3
	2 – Somewhat Accurate / Appropriate	2	2	1
	1 – Not at All Accurate / Appropriate	0	0	0
<b>Total</b>	4 – Extremely Accurate / Appropriate	279	270	272
	3 – Mostly Accurate / Appropriate	3	4	7
	2 – Somewhat Accurate / Appropriate	11	23	21
	1 – Not at All Accurate / Appropriate	7	3	0

*Table 3: Mathematics Teacher Review Ratings Grade 7*

Teacher Reviewer ID	Scale	Mathematical Language	Mathematical Vocabulary	Mathematical Content	Visual Representations	Distractors
1	4 – Extremely Appropriate	53	52	50	53	58
	3 – Mostly Appropriate	41	43	39	42	41
	2 – Somewhat Appropriate	5	4	11	5	0
	1 – Not at all Appropriate	1	1	0	0	1
2	4 – Extremely Appropriate	87	91	77	87	72
	3 – Mostly Appropriate	11	9	13	10	28
	2 – Somewhat Appropriate	2	0	5	2	0
	1 – Not at all Appropriate	0	0	5	1	0
3	4 – Extremely Appropriate	39	44	49	27	100
	3 – Mostly Appropriate	57	50	44	65	0
	2 – Somewhat Appropriate	3	4	3	5	0
	1 – Not at all Appropriate	1	2	4	3	0
Total	4 – Extremely Appropriate	179	187	176	167	230
	3 – Mostly Appropriate	109	102	96	117	69
	2 – Somewhat Appropriate	10	8	19	12	0
	1 – Not at all Appropriate	2	3	9	4	1

*Table 4: Mathematics Teacher Review Ratings for Bias Grade 7*

Teacher Reviewer ID	No Bias	Some Bias	Bias
1	95	5	0
2	95	2	3
3	100	0	0
Total	290	7	3

## Appendix G: Grade 8 External Review Data

*Table 1: Content Sampling Matrix Grade 8*

<b>Cognitive Engagement Level</b>	<b>Easy</b>	<b>Medium</b>	<b>Difficult</b>	<b>Total</b>
Procedural Fluency	25	34	36	95
Conceptual Understanding	46	44	20	110
Strategic Competence	18	19	16	53
Adaptive Reasoning	9	9	24	42
<b>Total</b>	<b>98</b>	<b>106</b>	<b>96</b>	<b>300</b>

*Table 2: Mathematics Education Expert Review Ratings Grade 8*

<b>Mathematics Education Expert ID</b>	<b>Scale</b>	<b>Mathematical Content</b>	<b>Mathematical Vocabulary</b>	<b>Distractors</b>
1*	4 – Extremely Accurate / Appropriate	97	96	98
	3 – Mostly Accurate / Appropriate	0	0	0
	2 – Somewhat Accurate / Appropriate	1	2	0
	1 – Not at All Accurate / Appropriate	2	2	2
2	4 – Extremely Accurate / Appropriate	91	90	81
	3 – Mostly Accurate / Appropriate	0	1	0
	2 – Somewhat Accurate / Appropriate	3	10	21
	1 – Not at All Accurate / Appropriate	8	1	0
3	4 – Extremely Accurate / Appropriate	84	91	93
	3 – Mostly Accurate / Appropriate	1	2	2
	2 – Somewhat Accurate / Appropriate	9	4	2
	1 – Not at All Accurate / Appropriate	3	0	0
<b>Total</b>	4 – Extremely Accurate / Appropriate	272	277	272
	3 – Mostly Accurate / Appropriate	1	3	2
	2 – Somewhat Accurate / Appropriate	13	16	23
	1 – Not at All Accurate / Appropriate	13	3	2

\*Mathematics Education Expert 1 did not review 1 of the 101 items in their set.



*Table 3: Mathematics Teacher Review Ratings Grade 8*

Teacher Reviewer ID	Scale	Mathematical Language	Mathematical Vocabulary	Mathematical Content	Visual Representations	Distractors
1	4 – Extremely Appropriate	71	78	76	78	77
	3 – Mostly Appropriate	28	18	19	18	22
	2 – Somewhat Appropriate	2	4	5	5	1
	1 – Not at all Appropriate	0	1	1	0	1
2	4 – Extremely Appropriate	82	86	78	88	62
	3 – Mostly Appropriate	12	15	17	5	35
	2 – Somewhat Appropriate	8	1	7	8	5
	1 – Not at all Appropriate	0	0	0	1	0
3	4 – Extremely Appropriate	35	40	43	36	95
	3 – Mostly Appropriate	61	55	53	59	1
	2 – Somewhat Appropriate	1	1	1	2	1
	1 – Not at all Appropriate	0	1	0	0	0
Total	4 – Extremely Appropriate	188	204	197	202	234
	3 – Mostly Appropriate	101	88	89	82	58
	2 – Somewhat Appropriate	11	6	13	15	7
	1 – Not at all Appropriate	0	2	1	1	1

*Table 4: Mathematics Teacher Review Ratings for Bias Grade 8*

Teacher Reviewer ID	No Bias	Some Bias	Bias
1	98	2	1
2	100	1	1
3	97	0	0
Total	295	3	2

---

## Appendix H: Item Writer Biographies

---

**Item Writer 1** holds a Doctoral degree in educational leadership with an emphasis in learning assessments and systems performance, Master's degrees in school psychology, special education, and counseling psychology, and a Bachelor's degree in psychology. She has experience as a program specialist and has co-designed district-wide mathematics. She is currently a school psychologist.

**Item Writer 2** holds a Master's of education in educational leadership, a Bachelor's of science in math education, and an ESL certification. She has 19 years of experience in education as a 7<sup>th</sup> and 8<sup>th</sup> grade mathematics teacher, mathematics curriculum specialist, ESC project coordinator for math and science, and in her current role as an administrative supervisor for math.

**Item Writer 3** Item Writer holds a Master's degree in educational administration, a Bachelor's degree in mathematics with a minor in psychology, and a credential in mathematics. She has 19 years of experience teaching Algebra II, Pre-Calculus, Advanced Placement Calculus AB, and Foundations of Mathematics/History of Mathematics. She is currently a mathematics facilitator.

**Item Writer 4** holds a Bachelor's and Master's degree in mathematics. She has 16 years of experience teaching all levels of algebra at the high school and college level. She is currently a graduate research assistant while pursuing her Doctoral degree in mathematics education.

**Item Writer 5** holds a Master's degree in education, a Bachelor's degree in interdisciplinary studies with a specialization in mathematics (1-8), and teaching credentials in mathematics and general education. She has 17 years of experience working in education as a mathematics teacher for Grades 6-8 and as an independent mathematics consultant.

**Item Writer 6** holds a Master's degree in education administration and a Bachelor's degree in mathematics with a minor in education. She has 13 years of experience teaching mathematics including Algebra I, Algebra II, and Geometry. She is currently an instructional facilitator for secondary mathematics.

**Item Writer 7** holds a Master's degree in education and Bachelor's degrees in sociology and education. She holds a general education certification for EC-4. She has 12 years of teaching experience and has taught students in Grades K-5. She has experience developing assessments and editing and reviewing district level curriculum.

**Item Writer 8** holds Master's degrees in special education and education policy and a Bachelor's degree in conflict resolution. She has 10 years of experience in education as a mathematics inclusion specialist, special education teacher, and in her current role as a policy and engagement fellow compiling and developing case studies about practices in education.

**Item Writer 9** holds Master's degrees in leadership of learning and secondary education, a Bachelor's degree in elementary education, and credentials in general education mathematics (1-8), and education administration (EC-12). She has taught 2<sup>nd</sup>, 6<sup>th</sup> and 8<sup>th</sup> grade mathematics. She is currently pursuing her Doctoral degree in curriculum and instruction.

**Item Writer 10** holds a Master's degree in leadership and policy studies and a Bachelor's degree in mathematics education. She has 8 years of experience in education as an academic coach, master math teacher, algebra teacher, elementary and middle school teacher, and in her current role as district mathematics supervisor.

**Item Writer 11** holds a Master's degree in educational administration, a Bachelor's degree in education, and credentials in general education (EC-4), gifted and talented, and English as a second language. She has 7 years of teaching experience in kindergarten and 4th grade.

**Item Writer 12** holds a Master's degree in childhood education and a Bachelor's degree in government. She has experience teaching elementary and middle school. She has also taught courses in teacher education and child development at the postsecondary level. She is currently pursuing her doctoral degree.

**Item Writer 13** holds a Bachelor's degree in finance, and a certification in education (4-8). He has 9 years of experience in education as a teacher, mathematics educator, instructional technology specialist, and in his current role as a mathematics department chair. He has experience developing and implementing professional development lessons for his mathematics department.

**Item Writer 14** holds a Bachelor's degree in recreation program management and a credential in education (K-4). She has 8 years of experience in education as a mathematics instructional leader. She currently teaches 4<sup>th</sup> grade mathematics and science.

**Item Writer 15** holds a Bachelor's degree in speech communications and holds a general education certification (4-8), a mathematics certification (4-12) and an ESL certification (4-8). She has 8 years of experience in education as a teacher, mathematics coach, and in her current role as a mathematics specialist.

---

## Appendix I: Mathematics Teacher Reviewer Biographies

---

**Mathematics Teacher Reviewer 1** holds a Doctoral degree in teaching leadership, a Master's degree in curriculum and instruction, and a Bachelor's degree in sociology. She has 6 years of experience as a 1<sup>st</sup> and 3<sup>rd</sup> grade teacher, a 6<sup>th</sup> grade math teacher, and a 5<sup>th</sup> grade gifted/talented teacher. She currently teaches 5<sup>th</sup> grade.

**Mathematics Teacher Reviewer 2** holds a Master's degree in education and a Bachelor's degree in elementary education. She has 19 years of experience teaching 4<sup>th</sup>-6<sup>th</sup> grade mathematics and was recently awarded teacher of the year. She currently teaches 4<sup>th</sup> grade mathematics and writing.

**Mathematics Teacher Reviewer 3** holds a Master's degree in education and a Bachelor's degree in interdisciplinary studies with a specialization in mathematics (Grades 1-8), She has 17 years of experience teaching 6<sup>th</sup> -8<sup>th</sup> grade mathematics.

**Mathematics Teacher Reviewer 4** holds a Master's and Bachelor's degree in mathematics. She has 16 years of experience teaching all levels of algebra at the high school and college level. She is currently a graduate research assistant while pursuing her Doctoral degree in mathematics education.

**Mathematics Teacher Reviewer # 5** holds a Master's degree in elementary education and a Bachelor's degree in elementary education. She has 17 years of experience teaching life skills and 3<sup>rd</sup> grade. She currently teaches 5<sup>th</sup> grade mathematics.

**Mathematics Teacher Reviewer 6** holds a Master's degree in public school administration and a Bachelor's degree in elementary education. She has 14 years of combined experience as a 6<sup>th</sup> – 8<sup>th</sup> grade mathematics teacher and as an Algebra I and Algebra II teacher. She currently teaches 6<sup>th</sup> and 7<sup>th</sup> grade mathematics.

**Mathematics Teacher Reviewer 7** holds a Master's degree in elementary education and a Bachelor's degree in interdisciplinary studies. She has 13 years of combined experience teaching Grades K-5 and as an ESL specialist. She currently teaches 2<sup>nd</sup> grade.

**Mathematics Teacher Reviewer 8** holds a Master's and Bachelor's degree in education and credentials in special education (K-12) and principal administration. He has 10 years of experience in education, teaching Pre-Algebra and Algebra I for 2 years. He is currently in his eighth year as a district math facilitator.

**Mathematics Teacher Reviewer 9** holds a Master's degree in education and Bachelor's degrees in sociology and education. She has 12 years of teaching experience and has taught Grades K-5. She has experience developing assessments and editing and reviewing district level curriculum.

**Mathematics Teacher Reviewer 10** holds a Master's degree in Education, Bachelor's degree in Elementary Education. She has 9 years of teaching experience as a 4<sup>th</sup> – 7<sup>th</sup> grade mathematics teacher. She currently teaches 4th grade mathematics and science.

**Mathematics Teacher Reviewer 11** holds a Master's degree in leadership and policy studies and a Bachelor's degree in mathematics education. She has 8 years of experience in education as an academic coach, master math teacher, high school algebra teacher, elementary and middle school teacher, and in her current role as district mathematics supervisor.

**Mathematics Teacher Reviewer 12** holds a Master's degree in education and a Bachelor's degree of art and science. She has 7 years of experience teaching 6<sup>th</sup> grade mathematics.

**Mathematics Teacher Reviewer 13** holds a Master's degree in curriculum and instruction, and a Bachelor's degree in elementary and early childhood education. She has 7 years of combined experience as a kindergarten and 4<sup>th</sup> grade teacher. She currently teaches 2<sup>nd</sup> grade.

**Mathematics Teacher Reviewer 14** holds a Bachelor's degree in interdisciplinary studies and a certification in ESL. She has 21 years of experience teaching Head Start, Kindergarten, 1<sup>st</sup> and 3<sup>rd</sup> grade. She currently teaches 3<sup>rd</sup> grade.

**Mathematics Teacher Reviewer 15** holds a Bachelor's degree in interdisciplinary studies. She has 11 years of teaching experience. She currently works as a team lead for 2<sup>nd</sup>, 3<sup>rd</sup>, and 5<sup>th</sup> grade.

**Mathematics Teacher Reviewer 16** holds a Bachelor's degree in Bilingual Education. She has 10 years of bilingual teaching experience. She currently works as a 2<sup>nd</sup> grade dual language teacher. She has experience serving on curriculum committees for Mathematics.

**Mathematics Teacher Reviewer 17** holds a Master's degree in curriculum and instruction and a Bachelor's degree in interdisciplinary studies. She has 6 years of experience teaching 4<sup>th</sup> and 5<sup>th</sup> grade mathematics. She currently teaches 5<sup>th</sup> grade.

**Mathematics Teacher Reviewer 18** holds a Master's degree in education and a Bachelor's degree in English. She has 6 years of combined teaching experience as a 4th and 5th grade bilingual teacher. She currently teaches in a 4<sup>th</sup> grade.

**Mathematics Teacher Reviewer 19** holds a Bachelor's degree in education (K-8) and is pursuing her Masters in math education. She has 7 years of experience teaching 1<sup>st</sup> – 3<sup>rd</sup> grade. She is currently a 2<sup>nd</sup> grade teacher.

**Mathematics Teacher Reviewer 20** holds a Bachelor's degree in recreation program management, and a credential in education (Grades K-4). She has 8 years of experience as a mathematics instructional leader. She currently teaches 4th grade mathematics and science.

**Mathematics Teacher Reviewer 21** holds a Bachelor's degree in interdisciplinary studies. He has 7 years of experience as a 6th grade mathematics teacher. He currently teaches 8<sup>th</sup> grade Pre-Advanced Placement Algebra I.

**Mathematics Teacher Reviewer 22** holds a Bachelor's degree in education and has 6 years of experience teaching Grades 1-5. She has also served as a member of a mathematics curriculum alignment. She currently teaches 3rd grade math and science.

**Mathematics Teacher Reviewer 23** holds a Bachelor's degree in applied arts and science. He has 3 years of experience as a Response to Intervention specialist and a 6th grade mathematics teacher.

---

## Appendix J: Mathematics Education Expert Reviewer Biographies

---

**Mathematics Education Expert Reviewer 1** holds a Doctoral degree in curriculum and instruction, a Master's degree second language education and culture, and a Bachelor's degree in education. She has 25 years of experience in education as a reading coordinator, tutor, ESL instructor, developmental writing teacher, mathematics content specialist, and instructional math coach. She currently works as a leadership development coach for Aspiring Leaders Teaching Trust.

**Mathematics Education Expert Reviewer 2** holds a Doctoral degree in mathematics, a Master's degree in mathematics, and a Bachelor's degree in mathematics. She has 14 years of experience in education as a mathematics instructor, teaching assistant, and tutor. She currently works as an assistant professor of mathematics researches mathematics curriculum and instruction.

**Mathematics Education Expert Reviewer 3** holds a Doctoral degree in mathematics education, a Master's degree in mathematics education, and a Bachelor's degree in mathematics. He has served as a teaching assistant and research assistant for 5 years at the university level. He is currently an assistant professor of mathematics education and serves as the principal investigator on a project that investigates student understanding of multivariate functions.

**Mathematics Education Expert Reviewer 4** holds a Master's degree in teaching and Bachelors' degrees in mathematics and general science. She has 13 years of experience in education as an algebra and biology teacher, professional development leader, and currently as a graduate student research assistant pursuing her Doctoral degree in mathematics education.

**Mathematics Education Expert Reviewer 5** holds a Master's degree in special education and a Bachelor's degree in history. She has 11 years of experience in education as a resource specialist (K-12), teaching assistant, assessment evaluator, program administrator intern. She has experience teaching special education students and developing and evaluating assessments. She is currently pursuing a Doctoral degree.

**Mathematics Education Expert Reviewer 6** holds a Master's degree in mathematics and a Bachelor's degree in mathematics and psychology. She has 7 years of experience in education teaching college level mathematics courses. She is currently an adjunct instructor of mathematics education.

**Mathematics Education Expert Reviewer 7** holds a Master's degree in mathematics and a Bachelor's degree in mathematics. He has 7 years of experience education teaching college Mathematics I and II, Linear Algebra, Differential Equations, Calculus and Multivariable Calculus. He is a member of a research team focused on geometry reasoning and instructional practices. He is currently a doctoral student.

**Mathematics Education Expert Reviewer 8** holds a Bachelor's degrees in psychology and English, and is currently working on her Doctoral degree in school psychology. She has 5 years of experience in education as a tutor, integrated practicum student, and special education para-educator. She currently works as a research assistant.



---

## **Appendix K – State Content Standards Referent Sources**

---

### **Texas**

The Texas Essential Knowledge and Skills (adoption 2012) were retrieved from:

<http://ritter.tea.state.tx.us/rules/tac/chapter111/index.html>

### **Common Core Standards**

The Common Core Standards in Mathematics were retrieved on June 10, 2011 from [www.corestandards.org/the-standards/mathematics](http://www.corestandards.org/the-standards/mathematics). These standards were published in 2010. They were developed as part of an initiative led by National Governors Association Center for Best Practices (NGA Center) and the Council of Chief State School Officers (CCSSO).

### **Virginia**

Virginia's Standards for Learning Document for Mathematics (adopted 2009 for full implementation in 2011-12) were retrieved from

[http://www.doe.virginia.gov/testing/sol/standards\\_docs/mathematics](http://www.doe.virginia.gov/testing/sol/standards_docs/mathematics)