SMU | ANNETTE CALDWELL SIMMONS SCHOOL OF EDUCATION & HUMAN DEVELOPMENT

RESEARCH IN MATHEMATICS EDUCATION

TIER Computations Progress Monitoring System: Expert Review

Technical Report 21-03

TIER Computations Progress Monitoring System: Expert Review

Anthony Sparks • Jennifer McMurrer • Leanne Ketterlin-Geller

Southern Methodist University

Spring 2021

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

This research was supported by the Texas Education Agency. Opinions expressed herein do not necessarily reflect those of the Texas Education Agency or individuals within.

Abstract

The purpose of the current report is to describe the external review process of item models for the Tiered Intervention with Evidenced-Based Research (TIER) Computations Progress Monitoring System. We recruited five external reviewers to critically analyze 220 item models that were developed by Texas teachers and assessment researchers. The results of these reviews helped inform the structure of the item models and the ensuing cousin item writing. This work is in collaboration with the University of Texas, Austin and the Texas Education Agency.

Table of Contents

TIER Computations Progress Monitoring System: Expert Review

Introduction

The purpose of the current report is to describe the expert review and reconciliation process of the item models for the progress monitoring instrument developed to measure computational fluency in grades K-6. We define essential vocabulary below:

- An **item** is a test question that is written to assess students' computational fluency, as described by the Texas Essential Knowledge and Skills (TEKS).
- An **item model** is a template that specifies the mathematical constraints for a specific item and is used in this project as the basis for future items (subsequently referred to as cousin items) assessing the same TEKS.
- A **cousin item** is an item written based on a specific item model. The primary goal for writing cousin items is to use the item model to create additional items that are parallel in difficulty, format, and response options.

The development of the item models is described in the Item Modeling technical report (Haider et al., 2021). The purpose of the external review is to seek expert feedback on the item models to support inferences about content validity. We sought feedback on whether the items aligned with the Texas Essential Knowledge and Skills (TEKS), whether the difficulty was appropriately specified, the ability of the item model to make 20 parallel items, and the appropriateness of the misconceptions. The review helped refine the item models before using them to write cousin items. The item models are used to create cousin items that are parallel in difficulty, format, and response options. Expert review provides content related evidence of validity to support the claim that the progress monitoring tool measures computational fluency as described by the TEKS (AERA, APA, & NCME, 2014)

Method

This section describes the expert review and reconciliation process for the item models.

Expert Review

We invited expert reviewers to review item models. Reviewers had to meet the following qualifications:

- a doctoral degree in mathematics, education, or related field;
- five years of experience working in a teaching, administrative, or university setting in their field;

- a deep understanding of the Texas Essential Knowledge and Skills;
- experience with writing mathematics assessment items in grades K-6; and
- extensive background in supporting elementary or middle school teachers as a mathematics coach, preferred.

Overall, we enlisted five expert reviewers to review the item models. Their bios are provided below. In the subsequent sections, we describe the external review process and results of the external review.

Reviewer 1 (Grade K) is an associate professor in school psychology at the University of Oregon. His research interest center on mathematical development, which includes assessment, instruction, and school systems that support mathematical development. Reviewer 1 has served or is currently serving as the principal investigator on twenty federally funded grants in mathematics instruction. He has also published articles and book chapters on mathematics instruction and assessment and developing multi-tiered instructional models. Prior to academia, Reviewer 1 was a practicing school psychologist.

Reviewer 2 (Grade 1) is an assistant professor and associate chair in the Department of Child and Adolescent Development at San Francisco State University. Her primary interest center on child development, early childhood education, early math and literacy development, professional development and teacher education, and the formation and implementation of early childhood public policy. Reviewer 2 has 14 years of experience working directly with children and families as a teacher and director of an early childhood program. She currently serves or has served in a consultant role for multiple organizations, including RTI International, UNESCO, and RUTU Foundation.

Reviewer 3 (Grades 2 & 3) is an assistant professor of special education at the University of Illinois, Urbana-Champaign. Her research interests include the teaching of mathematics to students experiencing difficulty, including students identified with disability labels. Currently, Reviewer 3 serves as the principal investigator for a project focused on pre-service teachers' views of disability and special education. She also received the Outstanding Dissertation Award from the American Education Research Association. Reviewer 3 actively publishes in special education and evaluation journals.

Reviewer 4 (Grades 4 & 5) is a postdoctoral researcher at Boise State University. She currently works on the National Science Foundation study titled "Research Order in Teaching (ROOT)". Reviewer 4 previously worked as a project coordinator on the Institute of Education Sciences grant titled "Recognizing Effective Special Education Teachers (RESET)". She also publishes extensively in the areas of observational tools and special education. Reviewer 4 has also held many teaching roles, including education specialist, middle school teacher, and elementary school teacher.

Reviewer 5 (Grade 6) is an assistant professor in the Department of Special Education and Communication Disorders at the University of Nebraska – Lincoln. Her research interest focus on the improving academic outcomes for students with learning disabilities. Reviewer 5 is

currently the principal investigator of an Institutes for Education Sciences subcontract grant "Exploring Cognitive and Foundational Processes Underlying Pre-algebra among Students with and without Mathematics Learning Disabilities". She is also an author and co-author of multiple interventions and assessments in the field of mathematics and special education.

We sent via email an excel file for external reviewers to input their feedback about the item models themselves. In the excel file, we asked expert reviewers to rate the item model on the following domains:

- Alignment with TEKS standards from blueprint;
- Alignment with level of difficulty assigned;
- Feasibility of item model constraints;
- Feasibility of item model producing 20 parallel forms;
- Appropriateness of the misconceptions (where applicable);
- Appropriateness of the alternate responses (where applicable); and
- Overall comments or other recommendations.

The rating scale was a four-point Likert scale (1: strongly disagree, 2: disagree; 3: agree; 4: strongly agree). After reviewing each item model within their respective grades, the expertl reviewers sent their feedback to the project manager.

Feedback Reconciliation

After receiving feedback from the expert reviewers, the RME team divided the process of reviewing the results and making changes to the item models. Each item model went through three reviews by the team. The first team member reviewed the feedback from the external reviewer and decided which aspects of the item model needed to be revised, if any. They made changes as necessary and documented the change and reasons for making the change in a shared excel tracking sheet. The second team member verified that the change was necessary, had been made on the item model template, and that the item model template was completely filled in. Lastly, the final team member reviewed the difficulty of the item model. This step was taken to analyze difficulty determinations across the models within a standard. The reconciliation team consisted of the following team members:

Leanne Ketterlin-Geller, Ph.D., is the principal investigator for the TIER Computations Progress Monitoring Tool project and Director at Research in Mathematics Education. Her research interests in include formative assessment design frameworks using modern test theory, including item response theory, empirical impact of accommodations and other test changes on the validity of test-score interpretations and uses, implications of using technology to implement universal design of assessment principles to support accessibility, and mathematics teachers'

decision-making with a focus on integrated research-based instructional design and delivery principles with content teacher knowledge.

Jennifer McMurrer, Ph.D., is the project manager for the TIER Computations Progress Monitoring Tool project and Senior Research Specialist at Research in Mathematics Education. She has worked in numerous research roles, including senior research analyst, and senior director of research. Dr. McMurrer's research interests focus on K-12 federal education policy and program implementation in persistently low-achieving schools.

M. Qadeer Haider, Ph.D., is a research and assessment coordinator at Research in Mathematics Education. Dr. Haider has developed and validated a reliable assessment tool to measure students' procedural and conceptual understanding of introductory linear algebra concepts as part of his doctoral dissertation. He also has mathematics and computer science teaching experience in Pakistan, Qatar, and the United States.

Tina Barton, M.A., is a design research strategist at Research in Mathematics Education. She currently works to implement Human-Centered Design principles and practices into the educational research and assessment development processes. She has teaching experience as a reading specialist and teacher.

Josh Geller, M.S., is a researcher at Research in Mathematics Education. He received his graduate degree from the University of Oregon in special education. Josh has also assisted on the creation of web-delivered math assessments researching the effectiveness of accommodations. He has also been a member of several item-writing teams.

Alain Mota, M.S., is a STEM Development and Implementation Coordinator at Research in Mathematics Education. Previously, he worked as a mathematics and science instructional coach for elementary and middle schools. Alain has also served as a science campus coordinator. He currently coaches teachers implementing intensive intervention in algebra readiness through "Project STAIR: Supporting Teachers of Algebra Individual Readiness".

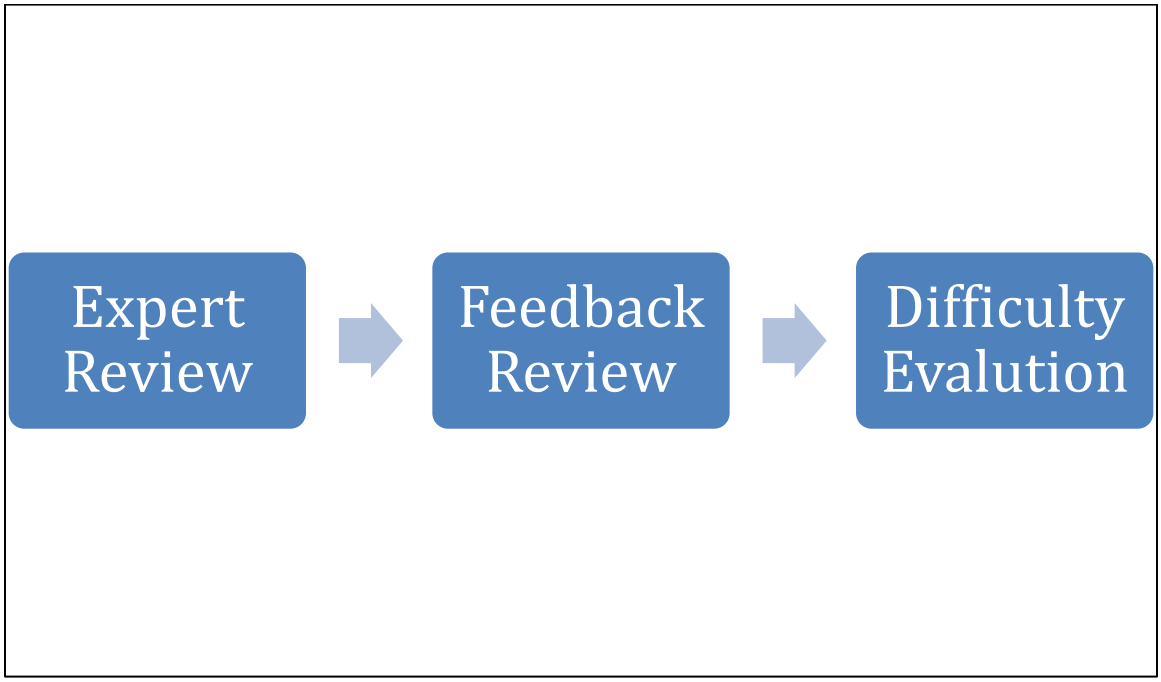
Robyn Pinilla, M.Ed., is a doctoral student at Southern Methodist University and a graduate research assistant at Research in Mathematics Education. Her research interests include problem solving capacity development in early childhood and elementary age children, and the impacts of mathematics instruction and intervention models on advancing adaptive reasoning capabilities. Robyn has experience as a special education teacher and administrator.

Results

In this section, we describe the results of the expert reviewers' feedback on the item models. Figure 1 explains our process for reviewing and reconciling the feedback. We begin by reviewing the experts' feedback. Then, we respond to the feedback by either updating the item model or noting why changes are not feasible. Lastly, we reevaluate the difficulty of the items to understand the distribution of item difficulty across item models. We describe the review of the feedback and reconciliation process in more detail below.

Figure 1

Review and Reconciliation Process



Feedback Review

Figure 2 highlights the feedback received from the reviewer in grade 4. If expert reviewers indicated a rating of disagree or strongly disagree, we asked the reviewer to provide feedback and/or recommendations for changes. Table 1 describes the percent agreement of the expert review across grades within the domains previously listed. The percent agreement of item models that aligned to the TEKS ranged from 77% to 100%. Alignment to the assigned difficulty agreement ranged from 57% to 95%. Agreement that the item constraints would yield 20 comparable items ranged from 77% to 100%. Agreement of the appropriateness of the misconceptions ranged from 44% to 100% and agreement in the appropriateness of the alternate responses ranged from 50% to 100%. The weakest domain across grades was the alternate responses with the highest domain being the misconceptions.

Figure 2

Sample Grade 4 Expert Review Feedback

Item Number	Primary Standard	Secondary Subtest	Difficulty Level (Easy, Medium, or Difficult)	Item model aligns with TEKS standard(s) from blueprint	Recommendations	Item model aligns with level of difficulty assigned	Recommendations
TIER_44A_01	**4(4)(A): Add and subtract whole numbers and decimals to the hundredths place using the standard algorithm.**	Add whole numbers and decimals (<1,000,000) - two whole numbers greater than 1,000 (within 20000 and 10000)	Easy	4		3	
TIER_44A_02		Add whole numbers and decimals (<1,000,000) - whole number and a decimal to the tenth or hundredths place (within 10000 and tenth place)	Easy	4		1	this could have regrouping errors, alignment errors, and additon with zero errors. Would that make it hard?
TIER_44A_03		Add whole numbers and decimals (<1,000,000) - decimals to the tenths place (both decimals to tenth place)	Medium	4		3	I am wondering if this is easy because the alignment of place values. But if an easy problem would involve no regrouping, I agree with this as medium.
TIER_44A_04		Add whole numbers and decimals (<1,000,000) - decimals to the hundredths place (both decimals to hundredths place)	Difficult	4		4	

Table 1

Expert Review Percent Agree/Strong Agree

Domain	K*	1*	2	3	4	5	6
Alignment to TEKS	100%	100%	77%	90%	93%	97%	100%
Difficulty Alignment	93%	95%	77%	90%	73%	57%	77%
Constraints	100%	98%	80%	100%	87%	77%	83%
Comparable Forms	0%	88%	100%	100%	83%	100%	100%
Misconceptions	44%	100%	100%	100%	100%	100%	100%
Alternate Responses	58%	100%	97%	97%	50%	50%	90%

*Note: Not all items had misconceptions/alternate responses

We highlight the review process by providing an example of feedback from one item model. This item model represents the TEKS 5(3)(B). The reviewer feedback on difficulty is provided in Figure 3. Out of a four-point scale, the reviewer rated the initial difficulty alignment as a two. The reviewer then made recommendations pertaining to the difficulty and how to constrain the item to match the designated difficulty.

Figure 3

Reviewer Feedback of 5(3)(B)_02

2	This seems to be medium as compared to 53B_01. Recommendation changing this to medium or constraining D or E to less than or equal to 2. At the same time it is easy compared to 53B_03

Figure 4 provides the response from the reconciliation team to TEKS 5(3)(B). An RME team member decided to constrain the item to keep its original difficulty designation. A second team member reviewed the response to the feedback and reconciled any differing opinions with the original team member.

Figure 4

Response to 5(3)(B)_02 Feedback

TIER_53B_02_KLK_01	additional constraint related to the value of the second multiplicand was added per reviewer's suggestion

Difficulty Evaluation

Figure 5 illustrates the last step in the reconciliation process. Two team members reviewed the difficulty of the item after changes were made during reconciliation. The first team member (center column) suggested switching the difficulty determination with a similar standard. The second team member (third column) agreed with the change and provided a justification.

Figure 5

Difficulty Determination for 5(3)(B)

TIER_5(3)(B)_02_KLK_01; TIER_5(3)(B)_03_KLK_01;	I see second item (TIER_5(3)(B)_03_KLK_01, marked Medium) easier than first item (TIER_5(3)(B)_02_KLK_01, marked Easy). I will recommend to swap their difficulty index.	I agree with this change due to the number of regroupings across the items.

TEA Review

TEA reviewed the combined item models within a template of a form for each of the seven grades (all items together within a grade) and provided a few comments on the item models after the RME team completed the reconciliation of the expert reviewers' feedback. Some comments pertained to the structure of the items themselves. For example, some graphics that were created for kindergarten needed adjustments to improve clarity of the item. Another comment for kindergarten centered on the prompt of the quantity discrimination section. For grade four, TEA made a comment about the lack of space for students to perform calculations. We were able to address the comments related to kindergarten by updating the graphics and the prompts. The design team at the University of Texas, Austin, formatted the assessment forms to allow for students to write their calculations.

Next Steps

Upon approval from TEA, the RME team resumed collaboration with the consultant educators to use the revised item models to write 20 cousin items for each of the 220 item models. These are items with the same difficulty, format, and response options. These items were used to populate 20 parallel forms for each of the seven grades (Kindergarten through grade 6). The cousin writing process is described in another technical report (Barton et al., 2021).

References

Barton, T., Sparks, A., McMurrer, J., & Ketterlin-Geller, L. (2021). *TIER computations progress monitoring system: Cousin writing* (Tech. Rep. No. 21-04). Dallas, TX: Southern Methodist University, Research in Mathematics Education.

Haider, M.Q., Sparks, A., McMurrer, J., & Ketterlin-Geller, L. (2021). *TIER computations progress monitoring system: Item modeling* (Tech. Rep. No. 21-02). Dallas, TX: Southern Methodist University, Research in Mathematics Education.