

Expanding a concept-specific measure of mathematical knowledge for teaching: Validation of the fractions instrument

Maryam Zolfaghari^{1*} , Christine Kathryn Austin² , Karl Wesley Kosko¹ , Hanan Alyami³ 

¹Kent State University, USA

²The University of Akron, USA

³Purdue University in Fort Wayne, USA

*Corresponding Author: mzolfagh@kent.edu

Citation: Zolfaghari, M., Austin, C. K., Kosko, K. W., & Alyami, H. (2026). Expanding a concept-specific measure of mathematical knowledge for teaching: Validation of the fractions instrument. *International Electronic Journal of Mathematics Education*, 21(2), em0871. <https://doi.org/10.29333/iejme/18096>

ARTICLE INFO

Received: 26 Aug 2025

Accepted: 22 Jan 2026

ABSTRACT

This study extends prior efforts to assess mathematical knowledge for teaching (MKT) by advancing the MKT-Fractions instrument, focusing specifically on teachers' knowledge for teaching fractions. Building an earlier measure centered on Knowledge of Content and Students (KCS), this phase integrates items targeting Specialized Content Knowledge (SCK) and Knowledge of Content and Teaching (KCT). Drawing on responses from 103 preservice teachers across two teacher preparation programs, we employed Rasch modeling to examine evidence for internal structure and item functioning. Results indicated strong item reliability, unidimensionality, and a coherent item hierarchy aligned with a construct map that represents the developmental trajectory of knowledge needed to teach fractions. The instrument effectively differentiated item difficulty across domains without significant variation in performance across programs. These findings support the use of the MKT-Fractions instrument as a tool for assessing preservice teachers' readiness and for informing teacher education practices focused on fraction instruction.

Keywords: mathematical knowledge for teaching, fraction, preservice teachers, Rasch modeling

INTRODUCTION

Teachers' ability to effectively teach fractions—a foundational yet challenging area of elementary mathematics—depends on more than procedural fluency or general content knowledge. Decades of research have shown that teaching mathematics well requires specialized knowledge that supports decision-making, error analysis, and instructional design in ways that go beyond solving problems correctly (Ball et al., 2008; Copur-Gencturk, 2021; Hill & Lubienski, 2007; Shulman, 1986). Shulman (1986, 1987) was among the first to formally articulate this distinction, proposing that teachers need both content knowledge (CK) and pedagogical content knowledge (PCK). His work underscored the unique intellectual demands of teaching, such as explaining mathematical ideas clearly, anticipating student misconceptions, and adapting instruction to support understanding.

Building on this foundation, Ball et al. (2008) introduced the framework of Mathematical Knowledge for Teaching (MKT) to more precisely identify the domains of knowledge required for teaching mathematics, which is defined as “as the mathematical knowledge needed to carry out the work of teaching mathematics” (Ball et al., 2008, p. 395).

Unlike the mathematics used in other professions, MKT includes the specialized reasoning and understanding required to support student learning, such as interpreting nonstandard strategies, selecting appropriate representations, and unpacking mathematical ideas during instruction. While many studies have examined MKT in broad terms, few have explored it through the lens of specific mathematical domains like fractions or distinguished among subdomains such as Specialized Content Knowledge (SCK), Knowledge of Content and Students (KCS), and Knowledge of Content and Teaching (KCT).

Recent efforts to develop and validate a fractions-specific MKT assessment have made progress in this direction. Prior iterations of the MKT-Fractions instrument (Kosko et al., 2019; Zolfaghari et al., 2021, 2024) focused on assessing KCS and provided strong validity evidence through test content, response processes, internal structure, and sensitivity to experience levels. Through these studies, we used cognitive interviews, construct mapping, Rasch modeling, and ANOVA comparisons to demonstrate that the instrument was both psychometrically sound and theoretically aligned.

This study represents the next phase in the progressive validation of the MKT-Fractions instrument. This study introduces a revised and expanded version of the assessment that includes new items targeting SCK and KCT, thereby broadening the scope of the instrument to better reflect the full range of MKT as described by Ball et al. (2008). Rather than reestablishing previous validity

claims, our aim in this study is to extend the existing validity argument by examining the internal structure and psychometric properties of the expanded instrument using Rasch modeling. Specifically, we investigate item and person fit, unidimensionality, and domain-level item difficulty, with particular attention to how empirical difficulty estimates align with the fractions-specific construct map, as well as relationships to participant variables.

BACKGROUND AND THEORETICAL FRAMEWORK

Fractions have consistently posed a conceptual challenge for both learners and teachers (Kazemi & Rafiepour, 2018; Ma & Ma, 1999). Misconceptions, such as treating numerators and denominators as separate whole numbers or overgeneralizing whole-number operations, are widespread and persistent. Addressing these misconceptions requires more than conceptual fluency; it demands that teachers possess the ability to interpret student thinking and adjust instruction accordingly (Copur-Gencturk, 2021; Tröbst et al., 2018). For instance, Copur-Gencturk (2021) found that many in-service teachers across the U.S. lacked a conceptual understanding of fraction operations, particularly division, limiting their ability to support student learning effectively.

Teaching fractions effectively, therefore, requires a kind of mathematical knowledge that is specific to the work of teaching. The Mathematical Knowledge for Teaching (MKT) framework (Ball et al., 2008) offers a comprehensive lens for identifying and organizing this knowledge. MKT distinguishes between content knowledge (CK) and pedagogical content knowledge (PCK), each with subdomains that reflect different aspects of teaching practice. Three subdomains are especially relevant for fraction instruction, and this paper includes: Specialized Content Knowledge (SCK), Knowledge of Content and Students (KCS), and Knowledge of Content and Teaching (KCT). SCK refers to mathematical knowledge unique to teaching. For example, evaluating the correctness of an unconventional strategy or analyzing a student's method independently of whether it was actually used in class. Importantly, SCK involves reasoning about mathematical ideas in ways that support instruction, but without necessarily drawing on student thinking or specific pedagogical moves (Copur-Gencturk et al., 2018). KCS emphasizes understanding how students typically think about specific content, including common misconceptions and developmental trajectories. KCT focuses on instructional decision-making, such as selecting appropriate representations, ordering examples, and adapting instructions to support students' conceptual understanding. For clarity, we consistently use the MKT framework terms SCK, KCS, and KCT throughout the manuscript to describe distinct forms of knowledge for teaching fractions, as illustrated in the examples that follow.

For instance, consider a student who incorrectly adds $1/4 + 1/3$ and writes $2/7$. To diagnose this error, a teacher would need SCK to recognize that the student is improperly adding numerators and denominators—an error that reflects a misapplication of whole-number addition procedures to fractions. This aligns with the idea that SCK involves identifying not only whether a solution is incorrect, but why it is mathematically flawed (Zolfaghari et al., 2024). Additionally, recognizing that this error reflects a common misconception among students (i.e., overgeneralizing whole-number addition) requires KCS. Teachers must anticipate that students often carry over their prior number sense when approaching fractions, leading to systematic misunderstandings (Ball et al., 2008).

Responding instructionally to such misconceptions and making pedagogically sound decisions about sequencing content or choosing representations—say, by using visual models like fraction bars or carefully sequenced task draws on KCT. For example, Tossavainen (2024) found that pre-service teachers frequently used visual aids such as pie models and everyday contexts (e.g., pizzas or cakes) to illustrate fractional concepts, particularly in early grades, and discussed the affordances and limitations of different representations. Similarly, Copur-Gencturk (2021) included prompts where teachers were asked to explain the use of common denominators or interpret unfamiliar denominators in division problems using visual or contextual examples, thereby capturing how teachers might employ pedagogical tools to deepen student understanding.

Although SCK, KCS, and KCT each reflect distinct aspects of the work of teaching, many existing MKT assessments fail to distinguish clearly between them—often combining items under broader categories like content knowledge (CK) or pedagogical content knowledge (PCK) (Copur-Gencturk et al., 2018; Depaepe et al., 2015; Zolfaghari et al., 2024). Moreover, few assessments account for how specific mathematical topics—such as fractions—demand varying levels and types of knowledge across these domains. The cognitive demands of SCK, KCS, and KCT tasks may differ significantly, yet little research has examined how this variation affects item difficulty or performance across populations. In response, scholars have called for concept-specific MKT measures and construct maps that capture the layered nature of teaching particular mathematical ideas (Jakobsen et al., 2014; Zolfaghari et al., 2021, 2024). These concerns about the specificity and complexity of MKT measures relate to ongoing discussions about how teachers' professional knowledge is structured. In particular, researchers have debated whether content knowledge and pedagogical content knowledge represent distinct but related constructs, or whether they are better captured as a unified dimension. In the next section, we examine how prior studies have approached this issue and what these findings suggest for the development of valid MKT assessments.

Existing Assessments of MKT and Its Dimensionality

These concerns about how different subdomains of MKT function in specific mathematical contexts raise broader questions about the structure of teacher knowledge. In particular, researchers have debated whether content knowledge and pedagogical content knowledge represent distinct yet related constructs, or whether they are best conceptualized along a single continuum. There has been considerable debate in recent years regarding the constructs used within and the dimensionality of teachers' professional knowledge for teaching mathematics. Central in this debate is whether mathematics content knowledge and pedagogical content knowledge assess the same underlying construct or if they are distinct, but correlated, constructs that assess teachers' professional knowledge (Blömeke et al., 2016; Charalambous et al., 2020; Copur-Gencturk et al., 2018). Rather, some

scholars have found that while content knowledge (CK) and pedagogical content knowledge (PCK) are strongly correlated, different factor analyses support a two-factor model (Blömeke et al., 2011, 2016; Hill et al., 2004; Kazemi & Rafiepour, 2018). Others have found evidence that both domains were best explained by a unidimensional, one factor model (Charalambous et al., 2020; Copur-Gencturk et al., 2018; Herbst & Kosko, 2014; Hill, 2010; Zolfaghari et al., 2024). Critical in these different analyses are how constructs have been defined (Charalambous et al., 2020; Copur-Gencturk et al., 2018), as the definitions of constructs and the conceptualization of how they interact directly affects the content of these items as they are written (Wilson, 2005).

Most scholars examining CK and PCK agree that CK includes knowledge of definitions, concepts, algorithms, and procedures whereas PCK includes knowledge about how to teach these concepts and procedures to students. Yet, even here, there are significant differences in the framing of these two domains. For example, Blömeke et al. (2016) define CK as including the skills to anticipate “typical student errors and on which parts of a student solution to focus” (p. 36). Thus, in their open-response example item for CK, participants were asked to solve two secondary math problems (presented in earlier secondary grades) in the first part of the item, and then to explain why Problem 2 would be more difficult for students than Problem 1. Ball et al. (2008) consider student errors and difficulties in two distinct ways. For example, SCK, as part of CK, involves “looking at patterns in student errors or in sizing up whether a nonstandard approach would work in general...” (p. 400). KCS, as part of PCK, focuses on anticipating students’ thinking, any potential errors, and why those errors may be enacted. Thus, SCK involves “sizing up the nature of an error...” KCS involves “...familiarity with common errors and deciding which of several errors students are most likely to make...” (p. 401). Depending on whether adopting the framework of Blömeke et al. (2016) or Ball et al. (2008), what is PCK for one measure is CK for another, or vice versa (Copur Gencturk et al., 2018).

Another potential issue is that many measures of professional knowledge include multiple mathematical concepts, and do not distinguish theoretically between what makes one item more or less difficult from another (Zolfaghari et al., 2021, 2024). To clarify, various scholars have found that PCK items tend to be more difficult than CK items (Copur-Gencturk et al., 2018; Depaepe et al., 2015; Tröbst et al., 2018), and there are a handful of studies that further distinguish between other MKT domains such that SCK items tend to be easier than KCT items, which are easier than KCS items (Herbst & Kosko, 2014; Ko et al., 2017). Yet there is a need for further specification for such results to be meaningfully applied to teacher education. Zolfaghari et al. (2021, 2024) argue for incorporating concept-specific frameworks that distinguish the work of teaching needed for specific topics. Specifically, Zolfaghari et al. (2021) found evidence that for PCK related to fractions, teachers’ ability to assess children’s creation and use of fractional parts is an easier teaching task than assessing children’s part-whole reasoning, which is easier than other more complex fraction tasks (i.e., fraction arithmetic). Evidence for the validity of this construct map (see **Figure 1**) has been presented across two different samples (Zolfaghari et al., 2021, 2024). In the current paper, we argue for not only the incorporation of such a construct map for levels of items, but an examination of the MKT domains (SCK, KCT, & KCS) in the context of such levels.

<p>Level 1 (SCK Example) Working with student(s) who can create and/or use of fractional parts</p> <p>Mr. Barnes asked Demetri to convert fractions to simplest form. Below are Demetri's answers to four fractions:</p> <p>A $\frac{4}{7} = \frac{2}{3}$ B $\frac{3}{7} = \frac{1}{3}$ C $\frac{3}{8} = \frac{1}{4}$ D $\frac{4}{8} = \frac{2}{4}$</p> <p>Based on what Mr. Barnes saw, what fraction should the expect Demetri to write when asked to simplify $\frac{1}{2}$?</p> <p><input checked="" type="radio"/> $\frac{1}{2}$</p> <p><input type="radio"/> $\frac{3}{4}$</p> <p><input type="radio"/> $\frac{1}{1}$</p> <p><input type="radio"/> $\frac{4}{3}$</p>	<p>Level 2 (KCT Example) Working with student(s) who can coordinate parts with the whole</p> <p>Mr. Diaz is beginning a unit on fraction comparisons with his 4th grade class. He knows many students are still developing part-whole reasoning. Given this, Mr. Diaz decides to begin having students compare $\frac{1}{3}$ and $\frac{1}{5}$.</p> <p>Of the options below, which should Mr. Diaz use to have students compare these fractions in the first lesson of the unit?</p> <p><input type="radio"/> Teach students the "butterfly method" for comparing fractions:</p> <p><input checked="" type="radio"/> Have students use fraction strips to help them draw and compare $\frac{1}{3}$ and $\frac{1}{5}$.</p> <p><input type="radio"/> Teach students to find the least common denominator to convert the fractions to $\frac{5}{15}$ and $\frac{3}{15}$.</p> <p><input type="radio"/> Each of the above options is equally valid.</p>
<p>Level 3 (KCT Example) Working with student(s) who can compare fractions & coordinate non-unit fractions</p> <p>Mr. Lawrence's 5th grade class is adding mixed numbers with unlike denominators. Mr. Lawrence walked by Miguel's desk and saw his work below:</p> <p>$4\frac{3}{8} = \frac{3}{8}$ $7\frac{1}{2} = \frac{7}{2}$</p> <p>$+ 3\frac{3}{8} = \frac{3}{8}$ $+ 4\frac{1}{2} = \frac{7}{2}$</p> <p>$\frac{3}{8} = 1\frac{1}{2}$ $\frac{7}{2} = 1\frac{1}{2}$</p> <p>Of the options below, which is the best course of action for Mr. Lawrence to help Miguel understand why his method is incorrect?</p> <p><input type="radio"/> Have Miguel correct both fractions in $4\frac{3}{8} + 3\frac{3}{8}$ to improper fractions ($\frac{35}{8} + \frac{27}{8}$) before finding a common denominator.</p> <p><input type="radio"/> Have Miguel draw a visual model to represent $4\frac{3}{8} + 3\frac{3}{8}$.</p> <p><input checked="" type="radio"/> Have Miguel draw a line to separate the whole numbers and the fractions:</p> <p><input type="radio"/> Have Miguel show why $\frac{3}{8}$ is equivalent to $\frac{6}{8}$ on a number line.</p>	<p>Level 4 (KCS Example) Working with student(s) who can coordinate fractions & wholes in more advanced circumstances (e.g., multiplying fractions)</p> <p>Mrs. Ashley asked Dan what fraction of the large rectangle is shaded? Dan look at the representation and said "it's one-half because they both are equal and one half of it shaded and half not"</p>  <p>Based on his response, which of the following describes Dan's reasoning?</p> <p><input type="radio"/> Dan said one-half because he cannot partition the square into equal parts.</p> <p><input type="radio"/> Dan said one-half because he can partition the square into equal parts, but doesn't understand these parts as fractions.</p> <p><input checked="" type="radio"/> Dan said one-half because it's a fraction he knows well and can visually recognize.</p> <p><input type="radio"/> Dan said one-half because he understands fractions as counting shaded parts and counting all parts</p>

Figure 1. Example MKT items across SCK, KCS, and KCT aligned with levels of the fractions construct map. Levels represent increasing instructional and cognitive demands associated with teaching fractions, ranging from creating and using fractional parts (Level 1) to coordinating fractions and wholes in more advanced contexts (Level 4) (Adapted from Kosko et al., 2022)

Overview and Purpose of the Study

In response to the limitations identified in prior research, this study builds on earlier work that focused exclusively on assessing Knowledge of Content and Students (Zolfaghari et al., 2021, 2024) by expanding the MKT-Fractions assessment to include new items targeting SCK and KCT. This expanded instrument enables a more comprehensive examination of how distinct types of MKT emerge in the context of teaching fractions, and it contributes to ongoing efforts to design valid, concept-specific tools for teacher education and research.

1. To what extent does the expanded MKT-Fractions measure, which now includes items targeting SCK and KCT, demonstrate evidence of internal structure validity through Rasch modeling?
2. How do items across the three MKT domains (SCK, KCS, and KCT) vary in difficulty, and what does this reveal about their dimensional coherence?
3. How does the revised MKT-Fractions measure perform across preservice teachers from two teacher preparation programs?

Instrument Development: MKT-Fraction

Prior versions of the MKT-Fractions measure focused exclusively on KCS items and collected validity evidence related to test content, response processes, reliability, and internal structure (Zolfaghari et al., 2021, 2024). In this study, we introduce an expanded version of the instrument that includes 20 multiple-choice items designed to assess three domains: KCS ($n = 10$), SCK ($n = 5$), and KCT ($n = 5$). Items were developed using evidence-centered design principles, with particular attention to domain-specific task features and explicit alignment with the construct map. As in previous studies, items are dichotomously scored and analyzed using Rasch modeling to examine internal structure and scale properties. It should be noted that one of the KCS items (F08) includes three embedded sub-items, each scored independently.

Figure 1 presents example items from each MKT domain, positioned according to their corresponding level on the construct map (Zolfaghari et al., 2024, p. 4). For example, in SCK03, teachers are given a student's pattern of fraction simplifications and asked to predict how the student might simplify $\frac{3}{4}$. This item, situated at Level 1 on the construct map, targets SCK by requiring recognition of procedural errors and reasoning patterns without explicit pedagogical framing. KCT item presents a scenario in which a teacher must select an appropriate method for comparing $\frac{1}{3}$ and $\frac{1}{5}$, taking into account students' developmental understanding of part-whole reasoning. This item targets KCT by requiring instructional decision-making tied to students' conceptual readiness and is positioned at Level 2 on the construct map, which involves working with students who can coordinate parts with the whole. The KCS item asks teachers to interpret a student's reasoning about a shaded rectangle, reflecting the need to anticipate a nuanced common misconception about equal partitioning; on the construct map, this appears at Level 4, where tasks require coordinating fractions and wholes in more advanced and nuanced situations.

METHOD AND ANALYSIS

Participants

This study involved 103 undergraduate students enrolled in elementary teacher preparation programs at two Midwestern U.S. universities located in different states (University A = 49.5%, University B = 50.5%). All participants had completed at least one mathematics methods course as part of their program requirements. Across both institutions, coursework emphasized foundational concepts in number, operations, and early algebraic thinking, with a shared focus on teaching fractions conceptually using visual models and attending to children's reasoning.

At University A, the teacher preparation program includes two consecutive 3-credit mathematics methods courses. These courses address number and operations, early algebra, geometry, fractions and decimals, and measurement. Field experience is embedded throughout both courses, and students' complete assignments that draw directly on their observations and teaching activities in elementary classrooms.

At University B, students are required to complete three 3-credit mathematics content courses focused on number and operations, probability and statistics, and algebraic and geometric thinking. One course is dedicated specifically to rational number sense and operations. In addition, students complete a single 3-credit mathematics methods course emphasizing pedagogical strategies for elementary mathematics instruction. This course includes a field component with required assignments.

Demographically, the sample consisted predominantly of female-identifying students (69.9%) with a small proportion identifying as male (6.8%) or transgender man (1%), with 22.3% of participants chose not to disclose their gender. Racial identities were self-reported using open text fields, resulting in diverse but inconsistently labeled entries. Most participants identified as White collectively representing approximately 90% of the sample. The remaining participants included students identifying as African American, Asian, Hispanic, and multiracial.

All participants completed the MKT-Fractions instrument in a single sitting, either online or on paper, as part of their university coursework. No compensation was provided. Data collection took place near the end of the academic term to ensure adequate exposure to methods course content and, for some, concurrent or recent field experience with upper elementary students.

Analysis

Building on prior phases of instrument validation (Zolfaghari et al., 2021, 2024), the present analysis evaluates the psychometric functioning of the expanded MKT-Fractions measure, which now includes items targeting SCK and KCT alongside previously established KCS items. Whereas earlier studies focused on collecting validity evidence for KCS-related content, this study broadens the scope to examine how the inclusion of additional MKT domains influences internal structure and domain-level performance.

Guided by the Standards for Educational and Psychological Testing (AERA et al., 2014), the analysis centers on two primary sources of validity evidence. First, the study examined internal structure validity using Rasch modeling to assess the degree to which item responses aligned with expectations of a unidimensional scale. This included analyses of item and person fit, separation indices, and reliability coefficients, as well as a Principal Component Analysis (PCA) of standardized residuals to detect potential multidimensionality. In addition, a Wright Map was constructed to visualize the alignment between item difficulty and participant ability across the three MKT domains. Second, to provide evidence based on relations to other variables, this study conducted an independent-samples t-test comparing participants from the two universities. This analysis explored whether total MKT-Fractions scores reflected meaningful variation across institutional contexts, thereby informing the generalizability and stability of the revised instrument.

Rasch modeling

Consistent with prior phases of instrument validation (Kosko et al., 2019; Zolfaghari et al., 2021, 2024), this study employed Rasch modeling to examine the internal structure of the expanded MKT-Fractions instrument. Since the nature of the assessment remained consistent—comprised of dichotomous, multiple-choice items—Rasch modeling was preferred for this phase of analysis. This model is especially well-suited for analyzing item response data of this kind, as it allows for more precise estimation of both item difficulty and person ability while transforming ordinal raw scores into interval-level data (Bond & Fox, 2015). Rasch analyses were conducted using Winsteps (Version 5.6.1.0).

Rasch modeling was used to assess multiple components of internal structure validity. To evaluate dimensionality, we conducted a Principal Component Analysis (PCA) of standardized residuals. This analysis is essential for determining whether the assessment measures a single latent construct or if there are indications of multiple underlying factors. PCA identifies contrasts in the residuals that may signal secondary dimensions, and we used eigenvalues, percentage of variance explained, and disattenuated correlations to interpret whether any observed contrasts reflected multidimensionality or a secondary strand within a unidimensional scale (Linacre, 2021). Additionally, we examined item and person fit statistics to evaluate the degree to which observed response patterns aligned with expectations under the Rasch model. Acceptable infit and outfit mean square (MNSQ) values, typically ranging from .75 to 1.33, indicate that items and persons fit the model well and that there is no substantial noise or misfit in the data (Bond & Fox, 2015; Linacre, 2021). In addition, item and person reliability indices were reviewed to assess the consistency and separation power of the measure. Item reliability reflects the degree to which item difficulty estimates are replicable across similar samples, while person reliability indicates the instrument's ability to distinguish among participants with varying levels of the latent trait (Linacre, 2021). Item and person separation indices were also examined to determine the number of statistically distinguishable levels of item difficulty and person ability present in the instrument.

Finally, Wright map was used to visually interpret the alignment between item difficulty and person ability on a shared logit scale. This mapping provides an intuitive representation of how items distribute across the continuum of knowledge, and it helps us examine how items across the three domains—Specialized Content Knowledge (SCK), Knowledge of Content and Students (KCS), and Knowledge of Content and Teaching (KCT)—cluster in terms of difficulty. The Wright Map serves as additional support for internal structure analysis and offers insight into the interpretive meaning of the scale.

T-test

In prior phases of validation, evidence based on relations to other variables has played a critical role in establishing the construct validity of the MKT-Fractions measure. Specifically, earlier studies examined how scores varied across levels of professional experience, such as comparing preservice and inservice teachers (Zolfaghari et al., 2021), as well as differences associated with field placement experience in upper elementary settings and enrollment status in teacher education programs (Zolfaghari et al., 2024). In a related line of inquiry, Kosko et al. (2025) demonstrated that teachers with higher PCK scores also exhibited more sophisticated professional noticing skills. These findings provided strong evidence that the instrument is sensitive to theoretically relevant group differences and aligned with broader indicators of teacher expertise.

In the present study, this approach is extended by using relations to other variables as a means of evaluating the generalizability of the revised instrument. Specifically, an independent-samples t-test was conducted to examine whether the expanded MKT-Fractions measure yields comparable results across preservice teachers from two teacher preparation programs located in different states. While both institutions serve similar populations and offer aligned coursework in foundational mathematics and pedagogy, the t-test allows us to assess whether the measure performs consistently across these distinct program contexts. This provides an additional layer of validation by testing the stability of instrument performance in two structurally similar, yet independently operating, teacher education environments.

RESULTS

To assess the validity of the MKT-Fractions measure, Rasch analysis, a well-established model for examining internal structure that accounts for both person ability and item difficulty (Bond & Fox, 2015), was employed. As a first step, this study conducted a

Table 1. Item analysis statistics

Item	Item difficulty	SE	Infit		Outfit		Point-biserial
			Mean square	Z	Mean square	Z	
F27	2.22	0.28	1.09	0.55	0.99	0.07	.21
KCT01	1.87	0.25	0.97	-0.18	1.12	0.59	.3
F16b	1.75	0.25	0.93	-0.07	0.92	-0.35	.33
F07	1.69	0.24	1.04	0.39	1.21	1.08	.23
F42	1.10	0.22	1.23	2.57	1.33	2.32	.09
F21	1.01	0.22	1.12	1.47	1.17	1.32	.22
F41	.91	0.22	1.04	0.51	1.06	0.58	.31
KCT03	.90	0.22	1.08	0.71	1.18	1.55	.25
KCT04	.50	0.21	1.05	0.66	1.06	0.6	.32
KCT02	.27	0.21	0.89	-1.46	0.89	-1.2	.49
SCK04	-.05	0.22	0.87	-1.7	0.82	-1.81	.53
SCK01	-.10	0.22	0.96	-0.49	0.96	-0.52	.29
F12	-.54	0.23	1.21	1.93	1.22	1.45	.15
F08pre_2	-.60	0.23	0.85	-1.44	0.76	-1.73	.54
KCT05	-.79	0.24	1.14	0.96	1.03	0.53	.45
SCK03	-1.00	0.25	1.04	0.19	0.91	-0.58	.53
F08pre_1	-1.06	0.25	0.99	-0.12	0.77	-1.18	.42
F44	-1.33	0.27	0.99	-0.21	1.01	-0.02	.33
SCK05	-1.33	0.27	0.83	-1.12	0.72	-1.18	.50
F08pre_3	-1.74	0.30	0.99	-0.46	1.16	-1.14	.43
F23	-1.74	0.30	0.94	-0.55	0.94	-0.85	.35
SCK02	-1.93	0.32	1.04	0.23	0.98	0.07	.24

Principal Component Analysis (PCA) of standardized residuals to investigate the dimensionality of the measure. This analysis determines whether a single latent construct accounts for the majority of item variance or whether substantial unexplained contrasts suggest the presence of additional dimensions (Linacre, 2021).

The total raw variance explained by the Rasch model was 31.7%, including 10.9% attributed to person ability and 20.7% to item difficulty. The first contrast in the residuals had an eigenvalue of 2.33, accounting for 7.3% of the total unexplained variance. According to Linacre (2021), an eigenvalue above 2.0 or unexplained variance greater than 5% may indicate the emergence of a secondary dimension that warrants further inspection. Although our first contrast met both thresholds indicating potential multidimensionality, this study examined disattenuated correlations between person estimates on the items contributing to the first contrast and the rest of the measure. These correlations were positive and high, suggesting that the items in the first contrast likely represent a secondary strand rather than a separate dimension (Linacre, 2021). As Linacre notes, “the suspect cluster of items is probably a secondary strand in the content area” and such patterns are typical in educational assessments without necessarily violating the assumption of unidimensionality, particularly when supported by theoretical coherence (Linacre, 2021).

This finding supports the interpretation that the MKT-Fractions instrument assesses a single, coherent construct — mathematical knowledge for teaching fractions — despite the inclusion of items from multiple MKT domains. As Linacre (2021) explains, first contrasts of this magnitude, when accompanied by strong positive disattenuated correlations, typically indicate the presence of a secondary content strand rather than a distinct latent dimension. In this case, the observed contrast is best understood as reflecting domain-related variation (SCK, KCS, and KCT) along a shared proficiency continuum, not a breakdown of unidimensionality. Such patterns are common in educational assessments that intentionally sample related but non-identical task features within a bounded content domain (Linacre, 2021).

Beyond dimensionality, we examined the extent to which individual items and person responses conformed to the expectations of the Rasch model. During the analysis, we examined item fit statistics to ensure all items aligned with the expectations of the Rasch model. One item, F14 demonstrated substantial misfit. Specifically, F14 showed an infit (MNSQ = 1.34, Z = 4.40), and outfit with (MNSQ = 1.48, Z = 4.38), well outside the recommended range of 0.75 to 1.33 (Linacre, 2021). In addition, F14 had a negative point-measure correlation (-0.09), suggesting that participants who generally performed well on the test tended to get this item incorrect, further indicating misalignment with the underlying construct. **Table 1** presents item statistics for the final set of retained items used in the Rasch analysis.

Given these fit violations, this item was removed from the final Rasch analysis. This exclusion led to an improvement in overall model performance, particularly in person reliability, which increased to 0.60, with a person separation index of 1.22. This improvement suggests that the removal of this item enhanced the measure’s ability to distinguish among participant ability levels, especially important in our sample of relatively homogeneous preservice teachers. Following item refinement, the final Rasch model demonstrated strong alignment between item functioning and model expectations.

The average item infit was (MNSQ = 1.00, Z = 0.10), and the average outfit was (MNSQ = 0.96, Z = -0.10), both within the acceptable range of 0.75 to 1.33, indicating that items functioned as expected within the model (Linacre, 2021). Similarly, person-level fit statistics showed alignment with model assumptions, with average infit and outfit scores of (MNSQ = 1.00, Z = 0.00) and (MNSQ = 0.97, Z = 0.00) respectively. These values suggest that participants’ responses were consistent with Rasch model expectations.

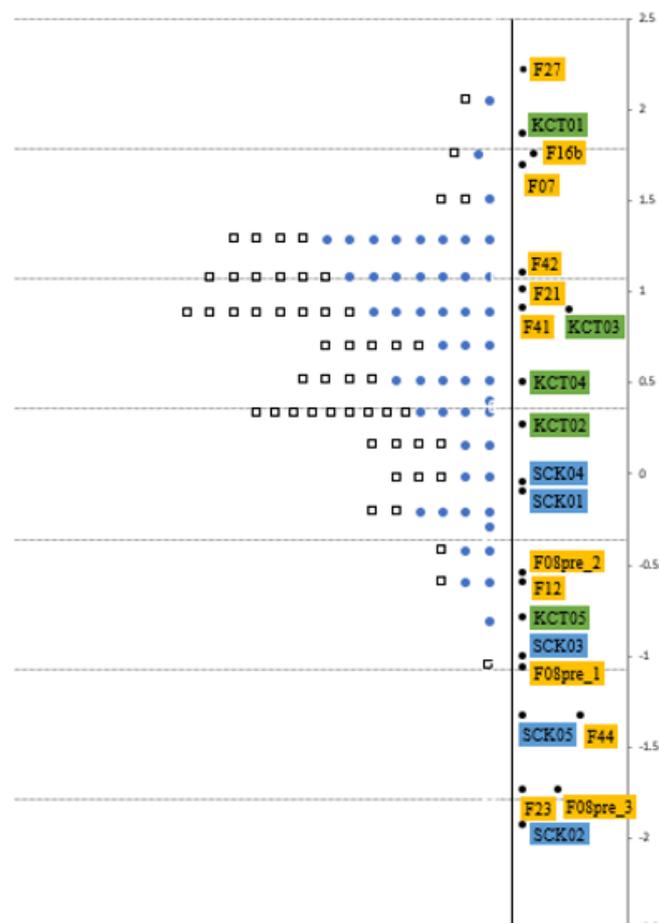


Figure 2. Wright Map displaying person ability (left) and item difficulty (right) on a shared logit scale, with items labeled by MKT domain (Source: Authors' own elaboration, using Winsteps (Version 5.6.1.0))

In terms of reliability, the instrument demonstrated high item reliability (.96), with a separation index of 4.90. This level of separation indicates that the measure is capable of distinguishing between at least five distinct levels of item difficulty, suggesting strong consistency in how items differentiate between lower and higher levels of mathematical knowledge for teaching fractions. Person reliability was more moderate (.60), with a separation participant experience and ability may reduce the model's ability to fully discriminate across ability levels (Kosko et al., 2019; Zolfaghari et al., 2021, 2024).

Alignment with Construct Map and Wright Map

Next, this study examined how item delta statistics, which signify their difficulty, aligned with our construct map for theorizing how items assessed different levels of MKT for fractions. An ordinal variable was created for each item signifying its level on the construct map (1= Level 1, 2 = Level 2, 3 = Level 3, 4 = Level 4) and calculated a Spearman Rho coefficient between construct map level and delta statistic for each item. Results suggest a strong and statistically significant positive association ($\rho = .67, p < .001$).

To further examine the internal structure of the MKT-Fractions measure, and theoretical alignment of the items, we analyzed the Wright Map. The Wright Map visually aligns person ability and item difficulty on the same logit scale. As illustrated in **Figure 2**, items assessing different MKT domains were distributed across a range of difficulty levels. As expected, Level 4 items (e.g., F27, KCT01, F16b) appeared toward the top of the scale, reflecting the complexity of tasks involving nuanced coordination of fractions and wholes, such as operations with uncommon denominators or fraction multiplication/division. Level 3 items (e.g., KCT03) occupied the upper-middle range, consistent with the coordination of non-unit fractions and comparative reasoning. Level 2 items (e.g., KCT02, SCK04) fell in the mid-range, representing part-whole coordination without full integration of equivalent fractions reasoning. Level 1 items (e.g., F23, SCK02) clustered toward the lower end, reflecting more foundational skills in creating and using fractional parts.

Overall, the correlation analysis and Wright Map patterns provide converging evidence that the empirical ordering of item difficulty aligns with the theoretical progression described in the construct map, while also highlighting a few individual items whose difficulty may warrant further review.

T-Test

To examine whether the MKT-Fractions instrument demonstrates stability across different institutional contexts, this study conducted an independent-samples t-test comparing preservice teachers' total scores from the two participating universities. The results of the independent t-test showed that there were no significant differences between preservice teachers at the two

universities with a negligible effect size, $t(101) = -0.014$, $p = .989$, $d = -0.03$. This finding suggests that participants performed similarly regardless of program affiliation, providing preliminary evidence for the generalizability of the instrument. Given that both universities offer comparable coursework in mathematics content and pedagogy, and that participants were at similar stages in their teacher preparation, this consistency in performance supports the claim that the revised MKT-Fractions measure functions reliably across comparable educational settings. As such, these results contribute to the validity argument based on relations to other variables by showing that the measure remains stable across similar but independent samples.

DISCUSSION

This study advances ongoing efforts to develop valid, concept-specific assessments of MKT by expanding the MKT-Fractions instrument beyond its original Knowledge of Content and Students (KCS) focus to include items targeting Specialized Content Knowledge (SCK) and Knowledge of Content and Teaching (KCT). While previous research has highlighted the persistent difficulty of teaching fractions (Ma & Ma, 1999; Siegler et al., 2011) and the need for assessments that distinguish between subdomains of MKT (Jakobsen et al., 2014; Zolfaghari et al., 2021, 2024). The use of construct map in this study, aligned with the Wright Map produced by our Rasch model, clearly articulates what teachers are capable of doing, pedagogically, with fractions when they have a specific score and which items assess their ability to do those specific things. Although a basic feature for validity evidence argued by Wilson (2005) and others, to our knowledge, no other MKT measure has employed this approach. By combining Rasch modeling with a construct map that specifies levels for fraction-related teaching tasks, our findings provide strong evidence that these domains can be meaningfully represented on a single dimension scale, empirically aligned with theoretical expectations, and applied consistently across different teacher preparation contexts. This integration of domain-specific content knowledge of fractions, informed pedagogical decision-making, and the ability to address teaching tasks of increasing complexity—aligned with the developmental trajectories of students' fraction understanding—positions the expanded MKT-Fractions instrument as a valuable tool for both research and practice in mathematics teacher education.

Building on prior validation work, the results of this study demonstrate that SCK, KCS, and KCT items can function along a shared measurement continuum when tightly bounded within the domain of fractions. Importantly, this unidimensional pattern reflects a shared proficiency continuum within the specific context of teaching fractions, rather than a collapse of theoretically distinct MKT domains. The unidimensional fit observed in the Rasch analysis suggests that variation in item features may reflect different cognitive and pedagogical demands rather than fundamentally separate constructs. For example, some SCK items were more difficult than some KCS items, but this had more to do with what they each assessed on the construct map than which MKT domain they were best aligned. The strong, positive correlation between construct map levels and empirical item difficulty further validates the hypothesized progression from foundational fraction tasks to more complex instructional decision-making. These findings align with and extend earlier research showing that MKT subdomains may be empirically intertwined when assessed within a specific mathematical topic.

The results also show that the majority of PSTs demonstrated above-average knowledge for teaching fractions ($M = 0.41$, $SD = 0.99$), where 0.00 represents the average level of knowledge. This outcome is unsurprising given that most participants had completed mathematics methods coursework and grades 3–6 field experience. The moderate level of person reliability likely reflects the combined effects of relatively high mean scores, a homogeneous preservice teacher sample, a smaller number of SCK and KCT items, and the absence of in-service teachers, which together limit the instrument's ability to fully discriminate across ability levels; future research that expands item difficulty and includes more diverse samples may strengthen person separation.

This work makes three primary contributions. First, it expands the MKT-Fractions instrument to capture the broader range of teacher knowledge needed for fraction instruction, moving beyond KCS to include SCK and KCT. Second, it integrates a construct map explicitly into the design and analysis process, allowing for empirical testing of theoretically grounded developmental progressions in fraction teaching tasks. Third, it provides evidence for the instrument's generalizability, with comparable performance across preservice teachers in two distinct preparation programs. Together, these contributions strengthen the validity argument for the instrument and position it as a resource for both research and teacher education.

The implications are twofold. For research, the findings contribute to ongoing debates about the dimensionality of teacher knowledge by showing that domain-specific CK and PCK subdomains can be meaningfully measured on a unified scale when the content is narrowly defined. This supports the feasibility of integrated measurement models in concept-specific contexts. For practice, the expanded instrument offers teacher educators a diagnostic tool to identify strengths and weaknesses across different aspects of fraction-related MKT. Such insights can inform curriculum design, professional development, and program evaluation by pinpointing which knowledge domains and task complexities require targeted support.

Several limitations should be noted. The sample was limited to preservice teachers from two Midwestern universities, restricting the generalizability of findings to broader populations and teaching contexts. The relative size of the SCK and KCT item pools was smaller than that of KCS, warranting further item development to ensure comprehensive domain coverage. Future research could extend this work by testing the instrument with in-service teachers, exploring longitudinal changes in MKT for fractions over the course of teacher preparation and early career teaching, and expanding the item pool to capture additional nuances in fraction instruction. Future validation work with larger samples will also examine differential item functioning across programs, experience levels, and MKT domains to further evaluate measurement invariance. Doing so would not only refine the measurement of fraction-specific MKT but also contribute to broader efforts to design robust, topic-focused assessments that bridge research and practice in mathematics education.

Author contributions: MZ, CA, & KK: conceptualization, methodology, instrument development, formal analysis, writing – original draft, writing - review & editing; **HA:** data collection, data curation, writing – review & editing. All authors have read and agreed to the published version of the manuscript.

Acknowledgements: Research reported here received support from the National Science Foundation (NSF) through DRK-12 Grant (1908159). Any opinions, findings, and conclusions or recommendations expressed in this article are those of the authors and do not necessarily reflect the views of National Science Foundation.

Funding: No funding source is reported for this study.

Ethical statement: The authors stated that the study was approved by Kent State University IRB (#19-232). All participants provided informed consent.

AI statement: The authors stated that generative AI tools were used only to assist with minor rewording and language editing of select sentences. All conceptualization, analysis, and results are solely the responsibility of the authors.

Declaration of interest: No conflict of interest is declared by the authors.

Data sharing statement: Data supporting the findings and conclusions are available upon request from the corresponding author.

REFERENCES

- American Educational Research Association (AERA), American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. American Educational Research Association.
- Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education*, 59(5), 389-407. <https://doi.org/10.1177/0022487108324554>
- Blömeke, S., Busse, A., Kaiser, G., König, J., & Suhl, U. (2016). The relation between content-specific and general teacher knowledge and skills. *Teaching and Teacher Education*, 56, 35-46. <https://doi.org/10.1016/j.tate.2016.02.003>
- Blömeke, S., Houang, R. T., & Suhl, U. (2011). TEDS-M: Diagnosing teacher knowledge by applying multidimensional item response theory and multiple-group models. *IERI monograph series: Issues and methodologies in large-scale assessments* (Vol. 4, pp. 109-129).
- Bond, T. G., & Fox, C. M. (2015). *Applying the Rasch model: Fundamental measurement in the human sciences* (3rd ed.) Routledge.
- Charalambous, C. Y., Hill, H. C., Chin, M. J., & McGinn, D. (2020). Mathematical content knowledge and knowledge for teaching: Exploring their distinguishability and contribution to student learning. *Journal of Mathematics Teacher Education*, 23(6), 579-613. <https://doi.org/10.1007/s10857-019-09443-2>
- Copur-Gencturk, Y. (2021). Teachers' conceptual understanding of fraction operations: results from a national sample of elementary school teachers. *Educational Studies in Mathematics*, 107(3), 525-545. <https://doi.org/10.1007/s10649-021-10033-4>
- Copur-Gencturk, Y., Tolar, T., Jacobson, E., & Fan, W. (2018). An empirical study of the dimensionality of the mathematical knowledge for teaching construct. *Journal of Teacher Education*, 70(5), 485-497. <https://doi.org/10.1177/0022487118761860>
- Depaeppe, F., Torbeyns, J., Vermeersch, N., Janssens, D., Janssen, R., Kelchtermans, G., Verschaffel, L., & Van Dooren, W. (2015). Teachers' content and pedagogical content knowledge on rational numbers: A comparison of prospective elementary and lower secondary school teachers. *Teaching and Teacher Education*, 47, 82-92. <https://doi.org/10.1016/j.tate.2014.12.009>
- Herbst, P., & Kosko, K. (2014). Mathematical Knowledge for Teaching and its Specificity to High School Geometry Instruction. In J. J. Lo, K. Leatham, & L. Van Zoest (Eds.), *Research trends in mathematics teacher education. Research in mathematics education*. Springer, Cham. https://doi.org/10.1007/978-3-319-02562-9_2
- Hill, H. C. (2010). The nature and predictors of elementary teachers' mathematical knowledge for teaching. *Journal for Research in Mathematics Education*, 41(5), 513-545. <https://doi.org/10.5951/jresmetheduc.41.5.0513>
- Hill, H. C., & Lubienski, S. T. (2007). Teachers' mathematics knowledge for teaching and school context: A study of California teachers. *Educational Policy*, 21(5), 747-768. <https://doi.org/10.1177/0895904807307061>
- Hill, H. C., Schilling, S. G., & Ball, D. L. (2004). Developing measures of teachers' mathematics knowledge for teaching. *The Elementary School Journal*, 105(1), 11-30. <https://doi.org/10.1086/428763>
- Jakobsen, A., Ribeiro, C. M., & Mellone, M. (2014). Norwegian prospective teachers' MKT when interpreting pupils' productions on a fraction task. *Nordic Studies in Mathematics Education*, 19(3-4), 135-150. <https://doi.org/10.7146/nomad.v19i3-4.148652>
- Kazemi, F., & Rafiepour, A. (2018). Developing a scale to measure content knowledge and pedagogy content knowledge of in-service elementary teachers on fractions. *International Journal of Science and Mathematics Education*, 16, 737-757. <https://doi.org/10.1007/s10763-016-9792-0>
- Ko, I., Milewski, A., & Herbst, P. (2017). *How are pre-service teachers' educational experiences related to their mathematical knowledge for teaching geometry?* In *The annual meeting of the American Educational Research Association Annual Meeting*.
- Kosko, K. W., Ferdig, R. E., & Zolfaghari, M. (2019). Preservice teachers' noticing in the context of 360 video. In S. Otten, A. G. Candela, Z. de Araujo, C. Haines, & C. Munter (Eds.), *Proceedings of the 41st annual meeting of the North American Chapter for the Psychology of Mathematics Education* (pp. 1167-1171). St. Louis, MO: University of Missouri.
- Kosko, K. W., Zolfaghari, M., & Austin, C. K. (2022). *The pedagogical content knowledge for fractions (PCK-Fractions) measure: User guide for educational researchers and teacher educators*. Kent State University.

- Kosko, K. W., Zolfaghari, M., Ferdig, R. E., Gandolfi, E., Austin, C. K., Pavlic, C., & Guan, Q. (2025). Effects of perceptual and cognitive factors on preservice teachers' professional noticing: A grounded cognitive perspective of professional noticing. *ZDM*, 58, 87-101. <https://doi.org/10.1007/s11858-025-01734-8>
- Linacre, J. M. (2021). *A user's guide to winsteps: Rasch-model computer programs*. Winsteps. <https://www.winsteps.com>
- Ma, L., & Ma, L. (1999). *Knowing and teaching elementary mathematics: Teachers' understanding of fundamental mathematics in China and the United States*. Lawrence Erlbaum. <https://doi.org/10.4324/9781410602589>
- Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15(2), 4-14. <https://doi.org/10.3102/0013189X015002004>
- Shulman, L. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, 57(1), 1-23. <https://doi.org/10.17763/haer.57.1.j463w79r56455411>
- Siegler, R. S., Thompson, C. A., & Schneider, M. (2011). An integrated theory of whole number and fractions development. *Cognitive Psychology*, 62(4), 273-296. <https://doi.org/10.1016/j.cogpsych.2011.03.001>
- Tossavainen, A. (2024). Pedagogical content knowledge in prospective elementary teachers' descriptions of teaching and learning of fractions. *Scandinavian Journal of Educational Research*, 69(5), 958-972. <https://doi.org/10.1080/00313831.2024.2362928>
- Tröbst, S., Kleickmann, T., Heinze, A., Bernholt, A., Rink, R., & Kunter, M. (2018). Teacher knowledge experiment: Testing mechanisms underlying the formation of preservice elementary school teachers' pedagogical content knowledge concerning fractions and fractional arithmetic. *Journal of Educational Psychology*, 110(8), 1049-1065. <https://doi.org/10.1037/edu0000260>
- Wilson, M. (2005). *Constructing measures: An item response modeling approach*. Lawrence Erlbaum Associates.
- Zolfaghari, M., Austin, C. K., & Kosko, K. W. (2021). Exploring teachers' pedagogical content knowledge of teaching fractions. *Investigations in Mathematics Learning*, 13(3), 230-248. <https://doi.org/10.1080/19477503.2021.1963145>
- Zolfaghari, M., Kosko, K. W., & Austin, C. K. (2024). Toward a Better Understanding of the Nature of Pedagogical Content Knowledge for Fractions: The Role of Experience. *Investigations in Mathematics Learning*, 1-18. <https://doi.org/10.1080/19477503.2024.2350909>