ETS

Measuring the Power of Learning.™

Research Memorandum

ETS RM–16-10

Investigating the Relevance and Importance of Mathematical Content Knowledge Areas for Beginning Elementary School Teachers

Michelle P. Martin-Raugh

Clyde M. Reese

Heather Howell

Richard J. Tannenbaum

Jonathan H. Steinberg

Jun Xu

August 2016

ETS Research Memorandum Series

EIGNOR EXECUTIVE EDITOR

James Carlson
Principal Psychometrician

ASSOCIATE EDITORS

Beata Beigman Klebanov
Senior Research Scientist

Heather Buzick
Research Scientist

Brent Bridgeman
Distinguished Presidential Appointee

Keelan Evanini
Research Director

Marna Golub-Smith
Principal Psychometrician

Shelby Haberman
Distinguished Presidential Appointee

Anastassia Loukina
Research Scientist

Donald Powers
Managing Principal Research Scientist

Gautam Puhan
Principal Psychometrician

John Sabatini
Managing Principal Research Scientist

Matthias von Davier
Senior Research Director

Rebecca Zwick
Distinguished Presidential Appointee

PRODUCTION EDITORS

Kim Fryer
Manager, Editing Services

Ayleen Gontz
Senior Editor

Since its 1947 founding, ETS has conducted and disseminated scientific research to support its products and services, and to advance the measurement and education fields. In keeping with these goals, ETS is committed to making its research freely available to the professional community and to the general public. Published accounts of ETS research, including papers in the ETS Research Memorandum series, undergo a formal peer-review process by ETS staff to ensure that they meet established scientific and professional standards. All such ETS-conducted peer reviews are in addition to any reviews that outside organizations may provide as part of their own publication processes. Peer review notwithstanding, the positions expressed in the ETS Research Memorandum series and other published accounts of ETS research are those of the authors and not necessarily those of the Officers and Trustees of Educational Testing Service.

The Daniel Eignor Editorship is named in honor of Dr. Daniel R. Eignor, who from 2001 until 2011 served the Research and Development division as Editor for the ETS Research Report series. The Eignor Editorship has been created to recognize the pivotal leadership role that Dr. Eignor played in the research publication process at ETS.

Investigating the Relevance and Importance of Mathematical Content Knowledge Areas for Beginning Elementary School Teachers

Michelle P. Martin-Raugh, Clyde M. Reese, Heather Howell, Richard J. Tannenbaum,
Jonathan H. Steinberg, and Jun Xu
Educational Testing Service, Princeton, New Jersey

August 2016

Corresponding author: Michelle P. Martin-Raugh, E-mail: mmartin-raugh@ets.org

Suggested citation: Martin-Raugh, M. P., Reese, C. M., Howell, H., Tannenbaum, R. J., Steinberg, J. H., and Xu, J. (2016). *Investigating the relevance and importance of mathematical content knowledge areas for beginning elementary school teachers* (Research Memorandum No. RM-16-10). Princeton, NJ: Educational Testing Service.

Find other ETS-published reports by searching the ETS ReSEARCHER database at http://search.ets.org/researcher/

To obtain a copy of an ETS research report, please visit http://www.ets.org/research/contact.html

Action Editor: Heather Buzick

Reviewers: Sam Rikoon and David Klieger

Copyright © 2016 by Educational Testing Service. All rights reserved.

ETS and the ETS logo are registered trademarks of Educational Testing Service (ETS).

MEASURING THE POWER OF LEARNING is a trademark of ETS. All other trademarks are the property of their respective owners.

Abstract

The purpose of this report is to explore the content-related validity evidence supporting the mathematics components of the *ETS*® National Observational Teaching Exam (NOTE) assessment series, a kindergarten through 6th grade teacher licensure assessment. To establish the content knowledge required for the effective teaching of mathematics in elementary school, we (a) identified content knowledge categories through the use of an expert panel and (b) surveyed a sample of 290 educators to verify that this body of content knowledge is indeed necessary and reasonable for the effective practice of beginning elementary school teachers teaching mathematics. We report information regarding the importance and relevance of mathematics content knowledge areas for both elementary school teachers and faculty members who prepare elementary school teachers. Implications of this work for the mathematics components of the NOTE assessment are discussed.

Key words: mathematics, content knowledge for teaching, elementary school, teacher licensure

Acknowledgments

Some of the content that appears in this report also is used in a companion report, entitled *Investigating the Relevance and Importance of English/Language Arts Content for Beginning Elementary School Teachers* (RM-16-08) by Michelle P. Martin-Raugh, Clyde M. Reese, Geoffrey C. Phelps, Richard J. Tannenbaum, Jonathan H. Steinberg, and Jun Xu.

Table of Contents

The purpose of this report is to explore the content-related validity evidence supporting the mathematics components of the *ETS*® National Observational Teaching Exam (NOTE) assessment series. NOTE is a kindergarten through 6th grade licensure assessment developed in a collaboration between Educational Testing Service (ETS) and TeachingWorks.[1] The NOTE assessment series is designed to measure a prospective elementary school teacher's ability to translate his or her content knowledge for teaching (CKT) into effective teaching practice. The NOTE assessment includes two components. One component includes standardized performance assessments that focus on three high-leverage teaching practices: modeling and explaining content, evaluating student thinking, and leading a classroom discussion. The second component focuses on the CKT used in teaching. Each component must include tasks that identify content that is necessary for the effective teaching of mathematics. A critical component of licensure assessments that focus on the beginning teaching proficiency in subjects such as mathematics are valid frameworks that define the mathematical content domains (American Educational Research Association [AERA], American Psychological Association [APA], National Council on Measurement in Education [NCME], 2014).

The remainder of this report is organized as follows. The first section provides a brief overview of the role of content-related validity evidence in licensure assessment. The second section describes the process used to generate the mathematical content framework. The third section describes the study design, methods, and results. Finally, we conclude by discussing the implications of this work for the mathematics components of the NOTE assessment series.

Content-Related Validity Evidence

Licensure is a legal requirement to enter or practice a profession, established by a governmental agency charged with overseeing the particular occupation or profession (Boulet & Zanten, 2014; Schmitt, 1995; Shimberg, 1981). Licensure assessments, as components of the licensure process, examine if candidates possess the knowledge and skills required for practice at the time of entry into the profession to help ensure the welfare of the public (Clauser, Margolis, & Case, 2006; Smith & Hambleton, 1990). Licensure assessments often measure the knowledge, skills, and abilities (KSAs) required for performing elements of a job rather than performance on actual job tasks (Wang, Schnipke, & Witt, 2005). Test specifications can describe assessment content and the KSAs that should be measured by the assessment (Raymond, 1996), and provide a critical foundation for validity evidence (Ebel & Frisbie, 1991). Consequently, it is essential to pinpoint the KSAs necessary for performing job tasks to design test specifications that are clearly related to performance in a given profession.

The *Standards for Educational and Psychological Testing* state that “validity is the most fundamental consideration in developing tests and evaluating tests” (AERA, APA, & NCME, 2014, p. 11). Strategies for validation that focus on content-related validity are emphasized in gathering evidence to support the use of licensure assessments (Raymond & Luecht, 2013; Shimberg, 1981). One job-analytic strategy that is often used to define the content domain for a licensure assessment involves using existing standards to compile a list of KSAs linked to the effective execution of job tasks and then confirm the KSAs using a survey of subject matter experts (SMEs; Rosenfeld & Tannenbaum, 1991; Tannenbaum & Wesley, 1993). The SMEs are selected based on their knowledge of and experience in the profession (Gael, 1983; Newman, Slaughter, & Taranath, 1999; Raymond, 2005; Raymond & Luecht, 2013; Rosenfeld & Tannenbaum, 1991). The survey contains of a list of KSAs necessary to perform the activities or responsibilities of the particular job effectively, and the SMEs are often asked to rate the KSAs regarding their relevance and importance (Kane, Kingsbury, Colton, & Estes, 1989; Raymond, 2005; Tannenbaum & Wesley, 1993). Surveys are efficient, as they allow for a large number of experts to provide information regarding a large number of knowledge or skill statements in an effective manner across multiple locations (Cascio, 1982; Raymond, 2001, 2005). Surveys also provide the opportunity to increase the representation and diversity of professional perspectives and try to ensure representation of minority subgroups (e.g., Black/African American, Hispanic/Latino) and differing geographic regions in the process of identifying KSAs most important for professional practice (Tannenbaum & Wesley, 1993). Other advantages include facilitating the development of assessment specifications (Kane, 1997; Raymond, 2001, 2005; Tannenbaum, Robustelli, & Baron, 2008) and documenting that the domain measured by the assessment is job related (Thompson & Thompson, 1982). However, it should be noted that the survey approach to job analysis is not without limitations. Complex information may be difficult to gather using a structured survey methodology (Raymond & Luecht, 2013). Moreover, in contrast with another commonly used job analysis methodology, the critical incident technique (Flanagan, 1954), the survey approach does not produce critical incidents that may be repurposed to develop selection or performance appraisal tools.

High-Leverage Content for Mathematics

In identifying the high-leverage mathematical CKT necessary to effectively teach the subject at the elementary school level, categories of student-level mathematical content were identified. The term “high leverage” is taken from the work of Deborah Ball and her colleagues (Ball & Forzani, 2011) to convey the idea that certain topics, because these make up a relatively larger portion of the curriculum or because failure to understand those can significantly undermine a student’s future mathematical progress, are critical for teachers to teach well.[2]

In an ideal world all teachers would be prepared on day one to teach all content effectively, but this is both unlikely and arguably not a reasonable expectation for beginning teachers given current systems of teacher preparation. The process of designating content as higher leverage takes into account that certain topics stand out as those that teachers really need to be prepared to teach. Fractions, for example, make up a large part of the student curriculum (National Governors Association [NGA] Center for Best Practices & Council of Chief State School Officers [CCSSO], 2010), are difficult for students to learn (Behr, Lesh, Post, & Silver, 1983; Carraher & Schliemann, 1991; National Assessment of Educational Progress [NAEP], 2005), are difficult for teachers to teach (Bruce & Ross, 2009), and failure to learn to work with fractions may undermine later student success in mathematics and in professional work (Reyna & Brainerd, 2008).

High-leverage content is “foundational to the ideas and skills of the K–12 curricula” (Ball & Forzani, 2011, p. 38). This statement is not meant to imply that other topics are not important, but it acknowledges that focusing both assessment and teacher education on these higher leverage topics is likely to have a disproportional payoff in terms of the effects on students.

In some ways the underlying idea behind high-leverage content is simple; some content matters more for students to know and it is more important that teachers be prepared to teach this content on day one. Identifying specific content as higher leverage via a rigorous process is, however, a complex task, and claims about the relative leverage of the content are based on a combination of curriculum analysis, analysis of the underlying mathematical structure and logical dependencies between ideas, and research on student learning. In the following section we describe the process by which high-leverage mathematical content was identified for the NOTE assessment series.

Initial Definition of the Domain

In identifying the mathematical content needed by beginning elementary school teachers, we cannot rely solely on identifying the mathematical knowledge and practices that students are expected to master and must additionally consider the mathematics teachers draw on in teaching that goes beyond that learned by students (Ball & Bass, 2000). A first step is to identify a subset of the student content domain that occupies a significant space within and across elementary grades and is fundamental for building future mathematical proficiency. A second step is to examine what mathematical knowledge is required in order to do the work of teaching this core, or high leverage, subset of student-level content, knowledge referred to in the field as CKT (Ball, Thames, & Phelps, 2008; ETS, 2011). Past efforts have found that while high-level descriptions of the work of teaching are useful frames, defining that work relative to the teaching of specific content is necessary to support the development of teacher assessments (Gitomer, Phelps, Weren, Howell, & Croft, 2014).

The domain-definition process was conducted by three ETS staff members, who were selected for their complementary expertise. One staff member is a former public school elementary mathematics teacher with 8 years of teaching experience and expertise in assessment design. The second is a researcher and former teacher educator with expertise in student learning of mathematics in grades pre-K to 5. The third is a researcher, former public school teacher, and former teacher educator with expertise in mathematics CKT and its assessment. This team began by drawing on the Common Core State Standards for Mathematics (CCSS-M; NGA & CCSSO, 2010) and other literature around student learning of the CCSS to create a list of mathematical topics covered by the CCSS in kindergarten through Grade 6, with a goal of covering the main content with a list of 10–15 distinct topic headings. Our goal was to build a category list of the content included in the CCSS in which categorized knowledge was finely enough delineated to make clear judgments about relative importance. Individual standards were classified by topic and possible overlaps or ambiguities identified and resolved via group consensus. The next step was for this three-person team to review this list of high-level topics against a set of filters adapted from guidelines developed at TeachingWorks and designed to support their ranking as higher or lower leverage content topics. Topics were evaluated against four guiding questions:

1. Is the content foundational (mathematically) to the ideas and skills of the K–12 curriculum?
2. Is the content taught in some form or another across several K–6 grade levels?
3. Does the content occupy a relatively large portion of the curriculum?
4. Is the content fundamental to students' learning in that it would constitute a source for students' difficulties when it is not well taught?

These questions, taken together, account for how much time students spend learning a particular topic, how much subsequent mathematical learning depends on the topic as a foundation, and how likely students are to struggle with the topic if the teacher is not well prepared to support students in learning it.

The team scored each topic against each guiding question using a three-point rubric—no, somewhat, and yes—and then met to reconcile scores. When consensus could not be reached among team members, the topic was brought to the larger development team for additional input. Table 1 summarizes the team's judgments for the 14 topics identified. (A description of the mathematical content areas is presented in the appendix.)

The outcomes then, of this initial domain-definition work were the following: (a) a list of 14 content topics covering the entire CCSS-M in Grades K–6 with accompanying descriptions and lists of associated standards and (b) a ranking of those topics with respect to their high-leverage status per the guiding questions adapted from TeachingWorks.

The CCSS-M also identifies mathematical practices that cut across grade levels and mathematical topics. Each of the eight mathematical practices identified by the authors of the CCSS-M were included on the content-related validity survey (described in the Survey subsection under Methods). No additional efforts were taken to reduce the list since the practices already represent those the field takes to be highest leverage. The appendix includes a description of the mathematical practices.

The Current Study

The current investigation surveyed educators to establish the importance of mathematical CKT for beginning elementary school teachers. It draws on a sample of practicing teachers and teacher educators to collect evidence of endorsement for the high-leverage content categories.

Table 1. Mathematics Content Knowledge for Teaching (CKT)

Mathematical topics	Ranking of topic according to high-leverage status	Guiding questions			
		1	2	3	4
CKT 5: Fractions and operations with fractions	8	Y	Y	Y	Y
CKT 9: Operations on whole numbers	8	Y	Y	Y	Y
CKT 10: Place value and decimals	8	Y	Y	Y	Y
CKT 3: Early equations and expressions	7	Y	Y	S	Y
CKT 7: Length, area, and volume	6	Y	Y	S	S
CKT 6: Integers and number lines	5	S	S	S	Y
CKT 2: Integers and number lines	4	Y	N	S	S
CKT 13: Shapes and angles	4	S	Y	S	N
CKT 11: Ratio and proportion and percents	3	S	N	N	Y
CKT 12: Rational and irrational numbers	3	S	S	S	N
CKT 1: Coordinate planes	2	S	N	N	S
CKT 4: Elementary data, variation, and distribution	2	S	N	S	N
CKT 8: Linear and simultaneous functions	2	S	S	N	N
CKT 14: Time and money	2	N	Y	N	N

Note. Rankings are of those topics with respect to their high-leverage status per the guiding questions adapted from TeachingWorks. The guiding questions are 1) Is the content foundational (mathematically) to the ideas and skills of the K–12 curriculum? 2) Is the content taught in some form or another across several K–6 grade levels? 3) Does the content occupy a relatively large portion of the curriculum? 4) Is the content fundamental to students' learning in that it would constitute a source for students' difficulties when it is not well taught? Y = yes; N = no; S = somewhat.

Methods[3]

Two versions of an online survey were constructed. Each version included two sections: the first asked respondents about high-leverage practices (common across content areas) and the second about high-leverage content (specific to a content area). The mathematics version couched judgments about the high-leverage practices in terms of teaching elementary school mathematics and included high-leverage content for teaching mathematics. Elementary school teachers and teacher preparation faculty were assigned to one of the two versions of the survey. The mathematics survey and the sample of educators responding to the mathematics survey are described in the Survey and Sample subsections.[4]

Survey

Relevance and importance judgments for each of 14 mathematical CKT areas and eight mathematical practices (MPs) for teaching mathematics were couched in a beginning teacher's ability to effectively teach the subject. For each mathematical content area or practice, two content-related validity questions were posed to educators:

1. Is knowing how to teach this content area **relevant** to a beginning elementary school teacher's ability to be an effective mathematics teacher?
2. If knowing how to teach this content area is relevant, how **important** is it to a beginning elementary school teacher's ability to be an effective mathematics teacher?

If educators indicated a CKT or MP as *relevant*, they then rated the importance of the CKT or MP (using a 6-point judgment scale). Therefore, importance ratings are only collected from respondents who judged the CKT or MP as relevant.

Following the relevance and importance judgments for the separate CKT areas and MPs, educators judged the three least and three most important content areas; an educator could not select the same CKT or MP for each category.

Sample

Working with a mailing list of 8,841 educators obtained from a national educational marketing firm, a multiphase outreach effort was conducted. The mailing list was sampled from a much larger, national database of teachers and teacher preparation faculty. The intent was to sample elementary school teachers and college faculty who prepare elementary school teachers. Sampling twice as many teachers as faculty members was intended to result in a significant number of teachers currently teaching lower (kindergarten to Grade 3) and upper (Grades 4 through 6) elementary classes as well as oversampling Black/African American and Hispanic/Latino teachers. The sample included an approximately equal number of teachers from each of the four United States Census regions. We also sought to oversample faculty from minority-serving institutions to better assure diversity in the sample and to allow for subgroup analyses, if samples permitted. Oversampling allowed us to gather data from a sample that roughly approximated the national distribution for teachers. Both public and private school teachers were included in the teacher sample.

Two versions of an online survey were constructed and educators were contacted, via email and letter, to invite them to complete one of the surveys. Participants were paid $25 in exchange for their participation in the study. Three rounds of e-mail follow-ups occurred during the data collection period to remind individuals to complete the survey. The assignment of teachers to one of the two versions of the survey was dependent on their current teaching assignments. Teachers who only taught mathematics were assigned to the mathematics version, and teachers who only taught English language arts (ELA) were assigned to the ELA version. Teachers who indicated they taught ELA and mathematics (more than 75% of the sample) or neither (approximately 4% of the sample) were randomly assigned to one of the two versions. Faculty also were randomly assigned to one of the two versions.

Of the original 8,841 educators contacted, 700 e-mails were not deliverable. Therefore, the number of educators successfully contacted was 8,141. Of these, 607 (or 7.5%) completed one of the two versions of the survey. An additional 31 educators were forwarded the survey by colleagues and completed it. In total, 638 educators completed either the mathematics or ELA version of the online survey.

Of the respondents, 387 (or 61%) indicated they were teachers (pre-K to Grade 12) and 202 (or 32%) indicated they were college faculty.[5] The remaining 49 (or 8%) respondents indicated they were administrators, held other education-related positions, or preferred not to provide information regarding their current position. Given the purpose of the survey, the 49 respondents who did not indicate they were teachers or faculty were removed from the sample.

The resulting sample—currently licensed teachers and college faculty currently preparing elementary school teacher candidates—includes 569 respondents, 385 teachers and 184 college faculty across the two versions of the survey.

While the overall response rate[6] for the survey was 7.5%, the resulting sample of elementary school teachers reasonably reflects the composition of the national population when compared to the National Center for Education Statistics (NCES) 2011–12 School and Staffing Survey (SASS; Goldring, Gray, & Bitterman, 2013), taking into account the sampling design considerations above. The sample of teachers slightly overrepresents the percent of Black/African American (11.5% in the sample compared to 7.1% nationally) and Hispanic/Latino (9.3% in the sample compared to 8.7% nationally) teachers compared to the latest SASS results (Goldring et al., 2013). The sample approximately mirrors elementary school teachers nationally in terms of years of experience, with approximately 40% of teachers with less than 10 years of experience and another 40% with more than 15 years of experience, and gender, with approximate 90% of elementary school teachers being female (Goldring et al., 2013).

The preceding discussion of response rates focused on the combined sample completing either the mathematics or ELA survey. For the following analyses, the sample is 290 educators who completed the mathematics version (188 teachers and 102 faculty). Table 2 provides a summary of background information for the mathematics sample overall. Tables 3 and 4 provide specific background information for teachers and faculty, respectively. Given the descriptive nature of the results presented for this study, the sample sizes of 200 to 400 respondents for the mathematics version can be viewed as adequate for generalizable findings (Kane, Miller, Trine, Becker, & Carson, 1995).

Table 2. Background Information—Overall Sample (*N* = 290)

Background information	*N*	%
Current position		
Teacher	188	65%
Faculty	102	35%
Gender		
Female	243	84%
Male	39	13%
Missing/prefer not to answer	8	3%
Race/ethnicity		
American Indian or Alaska Native	3	1%
Asian or Asian American	7	2%
Black/African American	35	12%
Hispanic/Latino	25	9%
White	196	68%
Two or more races	8	3%
Other/prefer not to answer/missing	16	6%
Geographic region		
Northeast	55	19%
Midwest	81	28%
South	97	33%
West	57	20%

Table 3. Teacher-Specific Background Information (*N* = 188)

Background information	*N*	%
Current teaching assignment		
Lower (Grades K–3)	108	57%
Upper (Grades 4–6)	59	31%
Other[a]	21	11%
Years of experience teaching elementary school		
3 years or less	26	14%
4 to 9 years	75	40%
10 to 14 years	36	19%
15 years or more	50	27%
Missing	1	‹1%
Mentored student teachers		
Yes	85	45%
No	102	54%
Missing	1	1%
Type of school		
Public (noncharter)	170	90%
Public (charter)	5	3%
Private	13	7%
School location		
Urban	68	36%
Suburban	79	42%
Rural	41	22%

[a]Teachers who taught across the elementary and secondary grades.

Table 4. Faculty-Specific Background Information (*N* = 102)

Background information	*N*	%
Years of experience		
3 years or less	9	9%
4 to 9 years	29	28%
10 to 14 years	19	19%
15 years or more	45	44%
Supervised student teachers		
Yes	81	79%
No	21	21%
Minority-serving institution		
Yes	27	26%
No	72	71%
Designation not available	3	3%
Institution location		
Urban	37	36%
Suburban	34	33%
Rural	31	30%

Analysis

There are two types of analyses in this report. The first type describes patterns in average relevance and importance judgment ratings in various ways, starting with all participants, and then making comparisons between teachers and faculty, lower elementary and upper elementary teachers, teachers across race/ethnicity groups, and teachers across geographic regions. Comparisons were made using effect sizes (Cohen, 1988) where the mean difference between two groups was divided by a combination of group sample sizes and standard deviations. In the case of race/ethnicity, White teachers were used as the reference group and in the case of geographic region, Northeastern teachers were used as the reference group. The second type of analysis indexes agreement between relevance and importance using intraclass correlations (ICC[2]; Shrout & Fleiss, 1979).

Results

The following results focus on the content-related validity evidence collected via an online survey from a sample of experienced educators—elementary school teachers and college faculty who prepare elementary school teachers. Results are reported for the overall sample (teachers and faculty).[7]

Relevance Judgments

Across the 14 CKT areas and eight MPs, the majority of educators, more than three quarters, agreed that all but three content areas and all practices are relevant for effective practice for beginning elementary school teachers teaching mathematics. The ICC[2] (Shrout & Fleiss, 1979) indexing agreement among educators regarding their relevance ratings across the 14 CKT areas and eight MPs is .98 (95% CI [.97, .99]), indicating near perfect agreement.

Three of the CKT areas received relevant judgments by less than 75% of teachers and/or faculty:

- CKT 1: Coordinate planes (59% of teachers and 71% of faculty);
- CKT 8: Linear and simultaneous functions (67% of teachers and 77% of faculty); and
- CKT 12: Rational and irrational numbers (59% of teachers and 77% of faculty).

The relevance judgments for teachers and faculty differed by 10 percentage points or more for the three mathematical content areas listed above as well as:

- CKT 11: Ratio and proportion and percentages (72% of teachers and 91% of faculty).

In all cases, the large differences in relevance judgments between teachers and faculty were the result of larger percentages for faculty.

Importance Judgments

Table 5 summarizes educators' judgments regarding the importance of each CKT and MP. The results for CKT 1 (Coordinate planes), CKT 8 (Linear and simultaneous functions), and CKT 12 (Rational and irrational numbers) should be interpreted with caution since less than 75% of the sample indicated the three CKT areas as relevant for beginning elementary school teachers and thus did not provide importance judgments. Results are presented for teachers, faculty and the total sample. The average importance judgment for all CKT areas and MPs, when combining teachers and faculty, was above 4.2 (on the six-point judgment scale)[8] in a range of 4.22 to 5.51 with variation across judgments from 0.71 to 1.07. Research by Tannenbaum and Rosenfeld (1994) recommended that an average importance judgment of 3.5 on a 5-point scale was sufficient to determine importance for licensure. Translating this finding to a 6-point scale would result in a threshold of 4.2. The differences in average importance judgments between teachers and faculty were lower than 0.25 (on the 6-point scale) for 12 of the 14 mathematical content areas and all 8 MPs. Effect sizes were lower than 0.20 (Cohen, 1988) for 12 of the 14 CKT areas and for 4 of the 8 MPs. Teachers and faculty differed most on:

- CKT 11: Ratio and proportion and percentages (4.42 by teachers and 4.74 by faculty; diff. = 0.32; ES = 0.32) and
- CKT 13: Shapes and angles (4.64 by teachers and 4.93 by faculty; diff = 0.29; ES = 0.33).

Table 5. Summary of Importance Judgments for Content Knowledge for Teaching (CKT) and Math Practices (MP)

Item	Teachers	Faculty	Overall
CKT 1	4.27 (1.11)	4.14 (1.00)	4.22 (1.07)
CKT 2	5.55 (0.67)	5.43 (0.78)	5.51 (0.71)
CKT 3	5.21 (0.77)	5.07 (0.83)	5.16 (0.79)
CKT 4	4.66 (0.89)	4.67 (0.79)	4.66 (0.85)
CKT 5	5.10 (0.83)	5.21 (0.74)	5.14 (0.80)
CKT 6	4.91 (0.92)	4.99 (0.93)	4.94 (0.93)
CKT 7	4.85 (0.77)	4.94 (0.88)	4.88 (0.81)
CKT 8	4.35 (0.96)	4.32 (0.98)	4.34 (0.97)
CKT 9	5.49 (0.76)	5.52 (0.69)	5.51 (0.73)
CKT 10	5.39 (0.74)	5.34 (0.87)	5.38 (0.79)
CKT 11	4.42 (0.98)	4.74 (0.99)	4.55 (0.99)
CKT 12	4.23 (1.02)	4.39 (1.01)	4.30 (1.01)
CKT 13	4.64 (0.90)	4.93 (0.83)	4.74 (0.88)
CKT 14	5.44 (0.74)	5.42 (0.84)	5.43 (0.77)
Minimum	4.23	4.14	4.22
Maximum	5.55	5.52	5.51
Sample Size	109–183	72–102	181–284
MP 1	5.20 (0.76)	5.20 (0.81)	5.20 (0.77)
MP 2	4.85 (0.93)	4.65 (0.98)	4.78 (0.95)
MP 3	4.66 (1.00)	4.56 (0.89)	4.63 (0.96)
MP 4	5.18 (0.89)	5.00 (0.87)	5.12 (0.88)
MP 5	5.11 (0.78)	5.10 (0.86)	5.11 (0.81)
MP 6	5.01 (0.93)	4.80 (0.92)	4.94 (0.93)
MP 7	4.76 (0.87)	4.56 (1.02)	4.69 (0.93)
MP 8	4.68 (0.90)	4.49 (1.05)	4.61 (0.96)
Minimum	4.66	4.49	4.61
Maximum	5.20	5.20	5.20
Sample Size	141–184	81–99	222–283

Note. Importance scale: 1 (*not at all important*), 2 (*of little importance*), 3 (*of some importance*), 4 (*moderately important*), 5 (*very important*), 6 (*extremely important*). Respondents who judged the practice not relevant are not included in the calculation of the average importance judgment.

Because respondents only made ratings for importance if they indicated a CKT or MP was relevant, this resulted in some missing cases. However, given the range in average importance judgments and associated levels of variation, there was no substantive evidence of floor or ceiling effects in the data. To compute ICC[2] we made the assumption that if a CKT or MP was not considered *relevant*, it would also be considered *not at all important*. Therefore, we imputed a value of 1 for missing cases. The ICC[2] (Shrout & Fleiss, 1979) indexing agreement among educators' regarding their importance ratings across the 14 CKT areas and eight mathematical practices is .99 (95% CI [.98, .99]).

In addition to considering the importance judgments of teachers overall, average judgments for teachers who are currently teaching lower (kindergarten through Grade 3) and upper (Grades 4 through 6) elementary grades were examined. Respondents to the survey were instructed to consider the full range of elementary grades when making their judgments; disaggregating teachers by current teaching assignment experiences revealed small differences, less than 0.25, in importance judgments for 10 of the 14 CKT areas and all of the MPs. Effect sizes were lower than 0.20 for seven of the 14 CKT areas and for seven of the eight MPs. For the remaining four areas, CKT 2 (Counting), CKT 3 (Early expressions and equations) and CKT 4 (Elementary data, variation, and distribution [higher judgments for lower grades]) and CKT 5 (Fractions and operations with fractions [higher judgments for upper grades]) had importance judgment differences greater than 0.25 with effect sizes of 0.41 or higher. Table 6 summarizes teachers' judgments, disaggregated by current grade levels taught.

Table 6. Summary of Importance Judgments for Content Knowledge for Teaching (CKT) and Math Practices (MP) by Current Grade Level Taught

Item	Lower (K–3)	Upper (4–6)	Difference
CKT 1	4.14 (1.20)	4.33 (1.04)	0.19 (0.17)
CKT 2	5.70 (0.55)	5.31 (0.80)	0.39 (0.60)
CKT 3	5.38 (0.69)	4.91 (0.86)	0.46 (0.62)
CKT 4	4.74 (0.91)	4.38 (0.81)	0.36 (0.41)
CKT 5	4.90 (0.89)	5.40 (0.65)	0.50 (0.61)
CKT 6	4.91 (0.97)	4.88 (0.86)	0.03 (0.03)
CKT 7	4.81 (0.86)	4.84 (0.60)	0.02 (0.03)
CKT 8	4.28 (0.97)	4.37 (1.02)	0.09 (0.09)
CKT 9	5.43 (0.82)	5.67 (0.54)	0.24 (0.33)
CKT 10	5.36 (0.77)	5.52 (0.75)	0.15 (0.20)
CKT 11	4.39 (0.96)	4.30 (0.96)	0.09 (0.09)
CKT 12	4.19 (1.00)	4.21 (1.14)	0.02 (0.02)
CKT 13	4.69 (0.85)	4.55 (0.98)	0.14 (0.15)
CKT 14	5.51 (0.75)	5.29 (0.73)	0.22 (0.30)
Minimum	4.14	4.21	0.02
Maximum	5.70	5.67	0.50
Sample Size	57-106	33-58	
MP 1	5.18 (0.82)	5.26 (0.61)	0.08 (0.11)
MP 2	4.85 (0.99)	4.91 (0.86)	0.07 (0.07)
MP 3	4.64 (0.98)	4.79 (1.00)	0.15 (0.15)
MP 4	5.17 (0.94)	5.22 (0.81)	0.05 (0.06)
MP 5	5.16 (0.84)	4.97 (0.70)	0.19 (0.25)
MP 6	5.02 (0.96)	5.05 (0.91)	0.03 (0.04)
MP 7	4.74 (0.88)	4.80 (0.87)	0.06 (0.06)
MP 8	4.61 (0.90)	4.76 (0.99)	0.15 (0.16)
Minimum	4.61	4.76	0.03
Maximum	5.18	5.26	0.19
Sample Size	78-106	45-58	

Note. Importance scale: 1 (*not at all important*), 2 (*of little importance*), 3 (*of some importance*), 4 (*moderately important*), 5 (*very important*), 6 (*extremely important*). Respondents who judged the practice not relevant are not included in the calculation of the average importance judgment.

Importance judgments broken down by educator ethnicity are shown in Table 7. Importance judgments for CKT areas by Black/African American respondents differed most from White respondents on CKT 1, CKT 4, CKT 8, CKT 11, CKT 12, and CKT 13 (higher means for Blacks/African Americans) and CKT 10 (Place value and decimals [higher mean for Whites]) with differences on these areas ranging from 0.31 to 0.97 on a 6-point scale with effect sizes ranging from 0.42 to 0.92. Two of these importance judgments by Hispanic/Latino respondents also differed most from White respondents (CKT 1 and CKT 12) at 0.28 (higher means for Hispanic/Latino respondents) on a 6-point scale with effect sizes of 0.25 and 0.27, respectively.

Table 7. Summary of Importance Judgments for Content Knowledge for Teaching (CKT) and Math Practices (MP) by Race/Ethnicity

Item	Black/African American	Hispanic/Latino	White	Overall
CKT 1	5.00 (0.71)	4.31 (1.11)	4.03 (1.11)	4.19 (1.10)
CKT 2	5.45 (0.67)	5.52 (0.67)	5.60 (0.67)	5.57 (0.67)
CKT 3	5.29 (0.64)	5.38 (0.74)	5.18 (0.81)	5.22 (0.78)
CKT 4	5.00 (0.75)	4.76 (1.00)	4.54 (0.87)	4.63 (0.88)
CKT 5	5.10 (0.70)	5.00 (0.95)	5.15 (0.83)	5.12 (0.82)
CKT 6	4.94 (0.80)	4.95 (0.91)	4.89 (0.95)	4.90 (0.92)
CKT 7	4.90 (0.72)	5.00 (0.87)	4.79 (0.73)	4.83 (0.75)
CKT 8	4.77 (0.73)	4.19 (0.98)	4.31 (0.97)	4.35 (0.95)
CKT 9	5.38 (0.67)	5.35 (1.03)	5.51 (0.75)	5.47 (0.78)
CKT 10	5.10 (0.91)	5.55 (0.80)	5.41 (0.67)	5.39 (0.73)
CKT 11	4.86 (0.77)	4.28 (1.02)	4.31 (0.97)	4.37 (0.96)
CKT 12	4.57 (0.85)	4.33 (0.90)	4.06 (1.05)	4.17 (1.02)
CKT 13	4.95 (0.79)	4.67 (1.06)	4.59 (0.88)	4.65 (0.90)
CKT 14	5.57 (0.60)	5.61 (0.72)	5.37 (0.77)	5.43 (0.75)
Minimum	4.57	4.19	4.03	4.17
Maximum	5.57	5.61	5.60	5.57
Sample Size	13–22	13–23	71–119	99–163
MP 1	5.19 (0.75)	5.05 (1.00)	5.24 (0.73)	5.21 (0.77)
MP 2	4.65 (1.11)	4.58 (1.02)	4.92 (0.90)	4.83 (0.95)
MP 3	4.69 (0.87)	4.38 (0.92)	4.77 (1.02)	4.70 (0.99)
MP 4	5.00 (1.05)	5.18 (0.96)	5.21 (0.87)	5.18 (0.90)
MP 5	5.38 (0.80)	5.09 (0.79)	5.08 (0.78)	5.12 (0.79)
MP 6	4.89 (0.94)	5.00 (1.15)	5.04 (0.89)	5.02 (0.93)
MP 7	4.88 (0.86)	4.68 (0.95)	4.76 (0.87)	4.77 (0.87)
MP 8	4.94 (0.87)	4.52 (0.87)	4.68 (0.95)	4.69 (0.93)
Minimum	4.65	4.38	4.68	4.69
Maximum	5.38	5.18	5.24	5.21
Sample Size	16–21	19–23	91–120	127–164

Note. Importance scale: 1 (*not at all important*), 2 (*of little importance*), 3 (*of some importance*), 4 (*moderately important*), 5 (*very important*), 6 (*extremely important*).

Among the MPs, importance judgments by Black/African American respondents differed most from White respondents on MP 2 (Reason abstractly and quantitatively [higher mean for Whites]), MP 5 (Use appropriate tools strategically), and MP 8 (Look for and express regularity in repeated reasoning [higher means for Blacks/African Americans]) with differences ranging from 0.27 to 0.31 and effect sizes ranging from 0.29 to 0.39. MP 2 also differed greatly for Hispanic/Latino respondents compared to White respondents (4.58 by Hispanics/Latinos and 4.92 by Whites; diff. = 0.34, ES = 0.37), but so did MP 3 (Construct viable arguments and critique the reasoning of others [4.38 by Hispanics/Latinos and 4.77 by Whites, diff. = 0.39, ES = 0.39]).

Importance judgments broken down by educator region are shown in Table 8. The differences in average importance judgments across regions (Northeast, Midwest, West, and South) were greatest on CKT 4, CKT 8, CKT 12, and CKT 13 with differences in these areas ranging from 0.41 to 0.56 on a 6-point scale. For CKT 4 and CKT 8, the average importance judgments for those from the Northeast were higher than those from the Midwest (ES = 0.45 and 0.27, respectively). For CKT 4 and CKT 12, those from the Northeast were higher than those from the West (ES = 0.28 and 0.38, respectively), but the reverse was true for CKT 8 (ES = 0.33). For CKT 13, judgments were higher in the South compared to those from the Northeast (ES = 0.44). Among the MPs, differences in average importance judgments across regions were greatest on MP 2, MP 3, and MP 5 in a range of 0.42 to 0.51 on a 6-point scale. For MP 2 and MP 3, the average importance judgment for those from the West was higher than those from the Northeast (ES = 0.48 and 0.36, respectively), while for MP 5, the average judgment for those from the Northeast was higher than those from the Midwest (ES = 0.39) and the West (ES = 0.38). It is also worth noting for MP 6 (Attend to precision), the average judgments for those from the South and the West were higher than those from the Northeast (ES = 0.32 and 0.24, respectively).

Table 8. Summary of Importance Judgments for Content Knowledge for Teaching (CKT) and Math Practices (MP) by Region

Item	Northeast	Midwest	South	West	Overall
CKT 1	4.44 (1.25)	4.09 (1.09)	4.31 (1.12)	4.35 (1.04)	4.27 (1.11)
CKT 2	5.51 (0.70)	5.52 (0.72)	5.63 (0.58)	5.50 (0.73)	5.55 (0.67)
CKT 3	5.09 (0.78)	5.13 (0.79)	5.31 (0.78)	5.31 (0.71)	5.21 (0.77)
CKT 4	4.80 (0.96)	4.40 (0.87)	4.85 (0.82)	4.54 (0.88)	4.66 (0.89)
CKT 5	5.21 (0.74)	4.98 (0.77)	5.08 (0.90)	5.25 (0.84)	5.10 (0.83)
CKT 6	4.96 (0.92)	4.67 (0.88)	5.05 (0.92)	5.00 (0.98)	4.91 (0.92)
CKT 7	4.78 (0.79)	4.77 (0.70)	4.90 (0.77)	4.96 (0.87)	4.85 (0.77)
CKT 8	4.42 (1.02)	4.16 (0.95)	4.32 (1.00)	4.71 (0.78)	4.35 (0.96)
CKT 9	5.45 (0.90)	5.48 (0.72)	5.50 (0.73)	5.55 (0.74)	5.49 (0.76)
CKT 10	5.24 (0.85)	5.32 (0.78)	5.52 (0.67)	5.45 (0.69)	5.39 (0.74)
CKT 11	4.50 (1.07)	4.28 (0.85)	4.41 (1.09)	4.61 (0.89)	4.42 (0.98)
CKT 12	4.41 (1.33)	4.23 (0.99)	4.30 (0.99)	4.00 (0.85)	4.23 (1.02)
CKT 13	4.54 (0.85)	4.35 (0.89)	4.91 (0.80)	4.66 (1.01)	4.64 (0.90)
CKT 14	5.44 (0.89)	5.44 (0.71)	5.45 (0.71)	5.39 (0.69)	5.44 (0.74)
Minimum	4.41	4.09	4.30	4.00	4.23
Maximum	5.51	5.52	5.63	5.55	5.55
Sample Size	17–35	31–55	36–66	20–30	109–183
MP 1	5.30 (0.73)	5.04 (0.81)	5.20 (0.72)	5.34 (0.77)	5.20 (0.76)
MP 2	4.76 (0.88)	4.66 (1.01)	4.92 (0.91)	5.17 (0.83)	4.85 (0.93)
MP 3	4.60 (1.00)	4.69 (0.85)	4.52 (1.08)	4.96 (1.04)	4.66 (1.00)
MP 4	5.25 (0.88)	5.00 (0.86)	5.22 (0.97)	5.34 (0.72)	5.18 (0.89)
MP 5	5.21 (0.69)	4.91 (0.81)	5.32 (0.68)	4.90 (0.92)	5.11 (0.78)
MP 6	4.84 (1.00)	4.92 (1.02)	5.13 (0.84)	5.07 (0.87)	5.01 (0.93)
MP 7	4.78 (0.75)	4.67 (0.90)	4.81 (0.87)	4.81 (0.94)	4.76 (0.87)
MP 8	4.60 (0.81)	4.66 (0.87)	4.67 (0.89)	4.81 (1.10)	4.68 (0.90)
Minimum	4.60	4.66	4.52	4.81	4.66
Maximum	5.30	5.04	5.32	5.34	5.20
Sample Size	25–34	44–54	49–66	23–30	141–184

In addition to judgments about importance at the individual CKT or MP level, educators were asked to identify the three most and three least important CKT areas and three most and three least important MPs. Table 9 summarizes the most and least importance judgments. Comparing relative importance judgments for teachers and faculty, the percentages differed by up to 29 percentage points for CKT areas and 19 percentage points for MPs. However, for the majority of CKT areas and MPs, the differences were less than 10 percentage points. For least important, three CKT areas and three MPs differed by more than 10 percentage points; for most important, four CKT areas and one MP differed by more than 10 percentage points.

Table 9. Summary of Least/Most Important Judgments for Content Knowledge for Teaching (CKT) and Math Practices (MP)

Item	Least Important		Most Important	
	Teachers	Faculty	Teachers	Faculty
CKT 1	86.2%	84.3%	1.6%	2.9%
CKT 2	6.4%	8.8%	83.5%	76.5%
CKT 3	12.2%	29.4%	59.0%	30.4%
CKT 4	48.9%	63.7%	12.2%	11.8%
CKT 5	16.5%	10.8%	44.7%	52.9%
CKT 6	30.3%	30.4%	23.9%	38.2%
CKT 7	22.3%	12.7%	23.9%	22.5%
CKT 8	82.4%	81.4%	2.1%	5.9%
CKT 9	3.7%	3.9%	79.3%	75.5%
CKT 10	5.9%	5.9%	75.0%	58.8%
CKT 11	75.0%	56.9%	6.4%	12.7%
CKT 12	83.0%	84.3%	1.6%	2.0%
CKT 13	21.8%	16.7%	21.3%	38.2%
CKT 14	5.3%	10.8%	65.4%	71.6%
Minimum	3.7%	3.9%	1.6%	2.0%
Maximum	86.2%	84.3%	83.5%	76.5%
MP 1	6.4%	2.9%	83.0%	87.3%
MP 2	49.5%	45.1%	22.9%	19.6%
MP 3	52.1%	57.8%	25.5%	28.4%
MP 4	8.5%	19.6%	73.4%	54.9%
MP 5	26.6%	13.7%	48.9%	50.0%
MP 6	37.2%	48.0%	25.0%	20.6%
MP 7	56.4%	51.0%	9.6%	18.6%
MP 8	63.3%	61.8%	11.7%	20.6%
Minimum	6.4%	2.9%	9.6%	18.6%
Maximum	63.3%	61.8%	83.0%	87.3%

To more easily digest the relative importance judgments, the *least* and *most* judgments were combined to rank order the CKT areas and MPs by perceived importance for beginning elementary school teachers. First, the CKT areas were ranked by the percentage of teachers who identified the CKT as one of the three most important; the highest percentage received a rank of 1 and the lowest a rank of 14. Second, the CKT areas were ranked by the percentage of teachers who identified the CKT as one of the three least important; the highest percentage received a rank of 14 and the lowest a rank of 1. Then the two rankings were summed with the lower values indicating the more relative importance. The process was repeated for the eight MPs (see Table 10).

Table 10. Rank Ordering of Content Knowledge for Teaching (CKT) Areas and Mathematical Practices (MP) by Relative Importance

Item	Teacher	Faculty
CKT 1	14	13
CKT 2	2	2
CKT 3	5	9
CKT 4	10	11
CKT 5	6	5
CKT 6	8 (T)	7 (T)
CKT 7	7	7 (T)
CKT 8	12	12
CKT 9	1	1
CKT 10	3 (T)	3
CKT 11	11	10
CKT 12	13	14
CKT 13	8 (T)	6
CKT 14	3 (T)	4
MP 1	1	1
MP 2	6	5 (T)
MP 3	5	5 (T)
MP 4	2	2 (T)
MP 5	3	2 (T)
MP 6	4	4
MP 7	7 (T)	8
MP 8	7 (T)	7

Note. (T) indicates a tie in the ranking of the relative importance.

The top four mathematical content areas—CKT 2 (Counting), CKT 9 (Operations on whole numbers), CKT 10 (Place value and decimals), and CKT 14 (Time and money)—were common between teachers and faculty. The top three mathematical practices—MP 1 (Make sense of problems and persevere in solving them), MP 4 (Model with mathematics), and MP 5 (Use appropriate tools strategically)—were common between teachers and faculty. [9]

All 14 CKT areas and eight MPs were generally judged to be important (average judgment about 4.2 or higher on a 6-point scale) by both teachers and faculty, and the relative importance indicates the two groups of educators have similar rankings of the more important content areas and practices.

Discussion

The purpose of the current investigation was two-fold—(a) examine the relevance and importance of mathematical content areas derived from the CCSS-M standards for students and (b) explore the content-related validity evidence supporting the test content specifications for the mathematics component in the NOTE assessment series. Content-validity evidence was gathered using an online survey of educators—practicing elementary school teachers and college faculty who prepare elementary teachers—who judged the relevance and importance of knowing how to teach each of the 14 CKT areas and eight MPs for beginning teachers. For 10 of the 14 CKT areas and all of the eight MPs, at least 75% of educators across both groups judged these to be relevant for beginning elementary school teachers. Three of the CKT areas judged as not relevant by more than a quarter of the educators—CKT 1, CKT 8, and CKT 12 — also were judged as the three least important CKT areas by more than 80% of the teachers and faculty. Each content area and practice was judged to be important, with six of the 14 content areas and three of the eight practices receiving an average judgment of 5.0 or higher (*very important*) on the 6-point scale by both teachers and faculty.

While there was less than complete agreement between teachers and faculty with respect to the most and least important content, when relative importance judgments were combined, both groups identified CKT 2, CKT 9, CKT 10, and CKT 14 as the top four content areas and MP 2, MP 6, and MP 7 as the top three practices. In other words, though both groups considered all of the listed content to be important, they were able, when faced with a forced choice, to classify some content areas as most important with relatively high levels of agreement.

Their classifications were also relatively consistent with our initial efforts to specify higher and lower leverage content, with one notable exception in the content category, CKT 14 (Time and Money), which was ranked quite low in our internal process and ranked fourth and third highest in importance by the faculty and teacher groups respectively. We believe that this points to a subtle difference between the criteria used in the two processes. The internal process evaluated mathematical content as *mathematics*, considering its importance with respect to student learning in an academic setting. Time and money was evaluated as not mathematically fundamental, as not comprising a large part of the curriculum, and as not undermining students' future mathematical learning if taught poorly, as there are limited connections between the ability to use time and money correctly and future mathematics that students are asked to engage in. Teachers, of course, must consider more than academics, and in making a more general judgment about importance, may have considered factors such as whether a topic is useful for life outside of school or whether it is a common idea that is likely to seem familiar to students, neither of which were considerations in our efforts to identify the highest leverage content. Time and money are, of course, commonly used representations in the curriculum, and also represent concepts that are important for students to master in order to become functional adults, although it may not be as mathematics that these topics are important in these ways.

One way to interpret the survey results is that, in the view of teachers and teacher educators, all of the content in the list of high-leverage content for teaching mathematics is sufficiently important to merit consideration in a licensure assessment for beginning elementary school teachers. Ideally, beginning and more experienced teachers would be fully prepared to teach all of the content of the student curriculum from day one. What is more difficult, both conceptually and practically, is identifying the relative importance of the content areas in order to focus our test specifications on the most important content for beginning teachers to be prepared to teach. The survey results confirm our assessment design in two important ways. First, these substantiate that even though teachers and teacher educators consider the full list of high-leverage content for teaching mathematics to be important, when asked to identify the most important content they are consistently able to do so. Second, with the one exception discussed previously, the external expert survey identified the same set of content topics as highest leverage as did our internal process, supporting the validity of favoring those particular topics in the assessment design.

Conclusion and Limitations

According to Sireci and Sukin (2013), the first step in developing an assessment is defining the domain being measured. In this investigation, we have collected evidence from practitioners in the form of relevance and importance judgments about the mathematical content and practices critical for effective practice for beginning elementary school teachers. The evidence confirms that the mathematical content areas and practices that are the focal aspects of the NOTE mathematics component are indeed relevant and important. This evidence, imperative for licensure assessment, may be used to create test specifications that characterize and operationally define the job domain (Raymond, 2001; Sireci & Sukin, 2013). Although the current study is an important first step in the assessment validation process, future research is warranted to examine the extent to which each test item is relevant to the job domain. Alignment studies, for example, may provide further evidence necessary to support the content validation of a licensure assessment such as NOTE.

The current research is not without limitations. First, although the structured survey methodology we used (cf., Raymond & Luecht, 2013) has many strengths that may outweigh its weaknesses, it produces simplified information that tends to be less rich than that obtained via focus groups or observations. Moreover, in contrast with other methodologies such as the critical incident technique (Flanagan, 1954), the survey approach does not produce critical incidents that may be repurposed to develop selection or performance appraisal tools. Second, although our sample consisted of educators from around the country, our sample was neither nationally representative of beginning teachers, practicing teachers, or teacher education faculty. Generalizations from the sample provided to educators in general or to subgroups of interest need to be made with caution. Finally, we also recognize that the mathematical content examined in this study includes a relatively large number of content dimensions and that the content defined under each dimension is itself extensive and complex. The data collected on agreement would be strengthened by evidence that participants understand the content dimensions as intended.

References

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

Ball, D. L., & Bass, H. (2000). Interweaving content and pedagogy in teaching and learning to teach: Knowing and using mathematics. In J. Boaler (Ed.), *Multiple perspectives on the teaching and learning of Mathematics* (pp. 83–104). Westport, CT: Ablex.

Ball, D. L., & Forzani, F. M. (2011). Building a common core for learning to teach: And connecting professional learning to practice. *American Educator, 35,* 17–38.

Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching. What makes it special? *Journal of Teacher Education, 59,* 389–407.

Behr, M. J., Lesh, R., Post, T. R., & Silver, E. A. (1983). Rational number concepts. In R. Lesh & M. Landau (Eds.), *Acquisition of mathematics concepts and processes* (pp. 91–126). New York, NY: Academic Press.

Boulet, J., & Zanten, M. (2014). Ensuring high-quality patient care: The role of accreditation, licensure, specialty certification and revalidation in medicine. *Medical Education, 48,* 75–86.

Bruce, C. D., & Ross, J. (2009). Conditions for effective use of interactive on-line learning objects: The case of a fractions computer-based learning sequence. *Electronic Journal of Mathematics and Technology, 3,* 12–29.

Carraher, D. W., & Schliemann, A. D. (1991). Children’s understanding of fractions as expressions of relative magnitude. In F. Furinghetti (Ed.), *Proceedings of the Fifteenth PME Conference* (Vol. 1, pp. 184–191). Retrieved from ERIC database. (ED413162)

Cascio, W. F. (1982). *Applied psychology in personnel management.* Reston, VA: Reston.

Clauser, B. E., Margolis, M. J., & Case, S. M. (2006). Testing for licensure and certification in the professions. *Educational Measurement, 4,* 701–731.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum.

Ebel, R. L., and Frisbie, D. A. (1991). *Essentials of Educational Measurement* (3rd ed.). Englewood Cliffs, NJ: Prentice Hall.

Educational Testing Service. (2011). *PRAXIS™ information brief.* Retrieved from https://www.ets.org/s/educator_licensure/ckt_handout.pdf

Flanagan, J. C. (1954). The critical incident technique. *Psychological Bulletin*, *51*, 327–358.

Gael, S. (1983). *Job analysis. A guide to assessing work activities.* San Francisco, CA: Jossey-Bass.

Gitomer, D. H., Phelps, G., Weren, B., Howell, H., & Croft, A. J. (2014). Evidence on the validity of content knowledge for teaching assessments. In T. J. Kane, K. A. Kerr, R. C. Pianta (Eds.), *Designing teacher evaluation systems: New guidance from the Measures of Effective Teaching project* (pp. 493–528). San Francisco, CA: Jossey-Bass.

Goldring, R., Gray, L., & Bitterman, A. (2013). *Characteristics of public and private elementary and secondary school teachers in the United States: Results from the 2011–12 Schools and Staffing Survey* (NCES 2013-314). Washington, DC: U.S. Department of Education, National Center for Education Statistics.

Kane, M. T. (1997). Model-based practice analysis and test specifications. *Applied Measurement in Education*, *10*, 5–18.

Kane, M. T., Kingsbury, C., Colton, D., & Estes, C. (1989). Combining data on criticality and frequency in developing plans for licensure and certification examinations. *Journal of Educational Measurement*, *26*, 17–27.

Kane, M. T., Miller, T., Trine, M., Becker, C., & Carson, K. (1995). The precision of practice analysis results in the professions. *Evaluation & the Health Professions*, *18*, 29–50.

Martin-Raugh, M. P., Reese, C. M., Phelps, G. C., Tannenbaum, R. J., Steinberg, J. H., & Xu, J. (2016). *Investigating the relevance and importance of high-leverage English language arts content knowledge areas for beginning elementary school teachers* (Research Memorandum No. RM-16-08). Princeton, NJ: Educational Testing Service.

National Assessment of Educational Progress. (2005). *The nation's report card: Mathematics 2005*. Retrieved from http://nces.ed.gov/nationsreportcard/pdf/main2005/2006453.pdf

National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common Core State Standards for Mathematics.* Washington, DC: Authors.

Newman, L. S., Slaughter, R. C., & Taranath, S. N. (1999, April). *The selection and use of rating scales in task surveys: A review of current job analysis practice.* Paper presented at the annual meeting of the National Council on Measurement in Education, Montreal, Canada.

Raymond, M. R. (1996). Establishing weights for test plans for licensure and certification examinations. *Applied Measurement in Education, 9*, 237–256.

Raymond, M. R. (2001). Job analysis and the specification of content for licensure and certification examinations. *Applied Measurement in Education, 14*, 369–415.

Raymond, M. R. (2005). An NCME instructional module on developing and administering practice analysis questionnaires. *Educational Measurement: Issues and Practice, 24*, 29–42.

Raymond, M. R., & Luecht, R. M. (2013). Licensure and certification testing. In K. Fl. Geisinger, B. A. Bracken, J. F. Carlson, J. I. C. Hansen, N. R. Kuncel, S. P. Reise, & M. C. Rodriguez (Eds.), *APA handbook of testing and assessment in psychology: Vol. 3. Testing and assessment in school psychology and education* (pp. 391–414). Washington, DC: American Psychological Association.

Reyna, V. F., & Brainerd, C. J. (2008). Numeracy, ratio bias, and denominator neglect in judgments of risk and probability. *Learning and Individual Differences, 18*(1), 89–107.

Rosenfeld, M., & Tannenbaum, R. J. (1991). *Identification of a core of important enabling skills for the NTE Successor Stage I Examination* (Research Report No. RR-91-37. Princeton, NJ: Educational Testing Service. http://dx.doi.org/10.1002/j.2333-8504.1991.tb01404.x

Schmitt, K. (1995). What is licensure? In J. C. Impara (Ed.), *Licensure testing: Purposes, procedures, and practices* (pp. 3–32). Lincoln, NE: Buros Institute of Mental Measurements.

Shimberg, B. (1981). Testing for licensure and certification. *American Psychologist, 36*, 1138–1146.

Shrout, P., & Fleiss, J. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86*, 420-428. doi:10.1037/0033-2909.86.2.420

Sireci, S., & Sukin, T. (2013). Test validity. In K. F. Geisinger & B. A. Bracken (Eds.), *APA Handbook of Testing and Assessment in Psychology: Vol. 1. Test theory and testing and assessment in industrial and organizational psychology* (pp. 61–84). Washington, DC: American Psychological Association.

Smith, I. L., & Hambleton, R. K. (1990). Content validity studies of licensing examinations. *Educational Measurement: Issues & Practice, 9*(4), 7–10.

Tannenbaum, R. J., Robustelli, S. L., & Baron, P. A. (2008). Evidence-centered design: A lens through which the process of job analysis may be focused to guide the development of knowledge-based test content specifications. *CLEAR Exam Review, 19,* 26–33.

Tannenbaum, R. J., & Rosenfeld, M. (1994). Job analysis for teacher competency testing: Identification of basic skills important for all entry-level teachers. *Educational and Psychological Measurement*, *54*(1), 199–211.

Tannenbaum, R. J., & Wesley, S. (1993). Agreement between committee-based and field-based job analyses: A study in the context of licensure testing. *Journal of Applied Psychology, 78*(6), 975–980.

Thompson, D. E., & Thompson, T. A. (1982). Court standards for job analysis in test validation. *Personnel Psychology, 35,* 865–874.

Wang, N., Schnipke, D., & Witt, E. A. (2005). Use of knowledge, skill, and ability statements in developing licensure and certification examinations. *Educational Measurement: Issues and Practice, 24,* 15–22.

Appendix. High-Leverage Mathematical Content Areas and Practices

Mathematical Content Areas

CKT 1. Coordinate planes

This topic area includes the use of the standard coordinate plane, graphing ordered pairs in all four quadrants, representing real-world and mathematical problems, and interpreting the values of coordinate values in the first quadrant mathematically and with respect to these contexts. It also includes awareness of certain relationships between coordinate points: understanding as reflections those that vary only in one coordinate and finding distance between such pairs using absolute value.

CKT 2. Counting

This topic area includes counting and “skip counting” integers between 0 and 1000, counting on, making connections to cardinality, understanding one-to-one correspondence between numbers and objects being counted, and identifying relationships between counting and the concept of larger and smaller numbers (i.e., that the next number in the counting sequence is one larger and that a number later in the counting sequence is larger).

CKT 3. Early equations and expressions

This topic area includes understanding what it means for algebraic objects to be considered equivalent, understanding how the equals sign is used to represent relational equivalence, and understanding that equations maintain their equivalence status under certain algebraic manipulations. It includes determining whether equations are true, determining missing values that would make them true, solving equations using the four operations, and observing the standard order of operations as well as the use of parentheses and distribution of multiplication over addition. In later elementary grades, it also includes the use of the less-than and greater-than relational symbols (<, >), and solving relational statements by substitution.

CKT 4. Elementary data, variation, and distribution

This topic area includes categorizing, organizing, and representing data, using picture and bar graphs with single-unit scales, and using line plots where horizontal scale is marked off in whole-number units or later in fractional units. It also includes recognizing statistical questions and understanding the concept of distribution and how it is described by the number of data points, center, spread, and overall shape. It also includes being able to select and calculate measures of central tendency and being able to interpret these and other attributes such as variation in the context of the shape of the graphed data or in the real-world context of the data.

CKT 5. Fractions and operations with fractions

This topic area includes having conceptual understandings of fractions, including understanding them as a part-whole relationship, as a sum of unit fractions, and fractions as numbers, as well as having the flexibility to move back and forth between these conceptualizations. It includes understanding both proper and improper fractions, the idea of equipartitioning as a building block for understanding fractions as part-whole relationships, understanding fraction equivalence, and being able to use strategies for comparing fractions. It includes performing operations on fractions such as addition, subtraction, multiplication, and division between fractions as well as between fractions and whole numbers. It also includes understanding and being able to execute strategies for performing the operations, understanding and being able to produce representations of the operations (i.e., scaling), building intuition about how the operations work (e.g., determining whether multiplying by a fraction makes something larger or smaller), and understanding common procedures for performing these operations.

CKT 6. Integers and number lines

This topic area includes understanding representations of whole numbers and integers, concepts of positive and negative numbers, and the concept of absolute value. It also includes using number lines as representations of fractions and decimals, measuring quantities on a number line diagram, and solving word problems that involve addition and subtraction of time intervals by representing the problem on a number line. The topic area includes interpreting statements of inequality in the context of a number line and displaying numerical data in plots on a number line.

CKT 7. Length, area, and volume

This topic area includes describing measurable attributes of objects and comparing two objects with a measurable attribute in common, choosing appropriate measurement tools, and taking measurement. It also includes calculating and estimating areas and volumes and making conversions between units.

CKT 8. Linear and simultaneous functions

This topic area includes generating a pattern to follow a rule, identifying its features, and identifying relationships between corresponding terms of ordered pairs built from two numerical patterns.

CKT 9. Operations on whole numbers

This topic area includes understanding representations of addition, subtraction, multiplication, and division (including representations involving manipulatives, diagrams, and mental images) and understanding representations of operations as number sentences. The topic area also includes solving problems involving the four operations alone or in combination and the use of the distributive property (including the use of mental strategies or the standard algorithms as appropriate), extending to word problems involving the representation and use of such operations.

CKT 10. Place value and decimals

This topic area includes having a conceptual understanding of what the digits in a number represent and being fluent in composing and decomposing numbers into groupings. In early grades, the content emphasizes the use of drawings or manipulatives. It includes having an understanding of why grouping and ungrouping are advantageous in the standard algorithms for performing operations on multi-digit numbers, including decimals, and having procedural fluency in using these processes and in rounding.

CKT 11. Ratio and proportion and percents

This topic area focuses on understanding the concept of ratios, unit rates, equivalent ratios, and percents. It includes using ratio language to describe proportional relationships, solving real-world problems where ratios are appropriate, using tables to compare ratios, thinking of percents as ratios, and solving problems involving finding the whole when given a part and the percent.

CKT 12. Rational and irrational numbers

This topic area focuses on ways of understanding rational numbers as numbers and as part of a number system. It includes having a conception of positive and negative rational numbers as points on a number line and understanding how this concept orders the numbers. It also includes using positive and negative numbers to represent quantities in real-world contexts, including solving real-world problems and identifying the meaning of 0 in context.

CKT 13. Shapes and angles

This topic area includes reasoning with shapes and their attributes, decomposing shapes into equal parts and using appropriate fraction terms to identify the parts, and drawing shapes based on specific attributes such as number of angles and number of equal faces. In later grades, the topic area includes working with lines, line segments, rays, and angles in two-dimensional figures as well as classifying such figures based on properties.

CKT 14. Time and money

This topic area includes being able to tell time and write time using analog and digital clocks and solving word problems that involve addition and subtraction of time intervals, including representing problems on a number line. In later grades, the topic area also includes solving word problems that require working with equations, analyzing the relationship between a dependent and an independent variable, and graphing ordered pairs of distance and time.

Mathematical Practices

MP 1. Make sense of problems and persevere in solving them

This practice involves approaching problems by looking at their meaning, analyzing the problem structure, making conjectures, and planning the solution process. It might include solving a simpler form of the problem first, comparing it to similar problems, or changing course if a strategy is not working or if a solution is not making sense.

MP 2. Reason abstractly and quantitatively

This practice involves using flexible thinking in moving back and forth between a problem context and its decontextualized mathematical description. This involves creating an abstraction of a problem so that it can be worked with in terms of the symbolic representation and includes the ability to pause at any point to consider the underlying problem space.

MP 3. Construct viable arguments and critique the reasoning of others

This practice involves making and exploring conjectures, constructing arguments, communicating with others, and deciding whether others' arguments make sense. It can involve the use of case-based logic and counterexamples. At the elementary level, this work is likely to be less formal and structured more around concrete referents.

MP 4. Model with mathematics

This practice involves using mathematics to solve authentic problems. This may involve simplifying the situation or modeling only part of a situation in which the mathematical model is more general. It also involves being flexible in choosing an appropriate way to model a given problem.

MP 5. Use appropriate tools strategically

This practice involves choosing appropriate tools. These tools include everything from manipulatives to paper and pencil to computer applications. This practice focuses on becoming familiar with available tools, considering which tools might be appropriate, and making informed decisions about which tool is most appropriate for the task.

MP 6. Attend to precision

This practice involves choosing language carefully, particularly when communicating mathematics with others. This often involves specifying definitions, defining symbols and units, and labeling. It also involves choosing and adjusting the language used in making claims or conjectures in order to describe ideas precisely.

MP7. Look for and make use of structure

This practice involves noticing patterns in problems or solution processes and being able to step back from a particular problem or context to consider its similarity to or difference from other problems or contexts.

MP8. Look for and express regularity in repeated reasoning

This practice involves giving attention to reasoning processes as objects in themselves, seeing patterns in such processes, and looking for efficiencies. These efficiencies might include taking shortcuts or using repeated processes. This requires that students attend to their solution processes as they solve problems.

Notes

[1] TeachingWorks is housed in the School of Education at the University of Michigan and focuses on the improvement of teacher preparation.

[2] We note here that mathematical practices as identified in the CCSS are addressed separately, as they already represent the field's effort to identify highest leverage topics.

[3] The data collection for mathematics content and practices for teaching was embedded in a larger content-related validity evidence data collection.

[4] Results for the survey of high-level ELA content is reported in Martin-Raugh et al. (2016).

[5] Of the 387 respondents who indicated they were teachers, all but two indicated they held a current license to teach in their state. These two respondents were removed from the sample. Of the 202 respondents who indicated they were faculty, 17 indicated they were not currently preparing elementary school teacher candidates. These respondents also were removed from the sample.

[6] Response rate for the e-mailed surveys delivered to sampled educators.

[7] The sample sizes for Black/African American (n = 35) and Hispanic/Latino (n = 25) respondents are relatively small, but whether the judgments of these subgroups support or run counter to the findings for the overall sample will be highlighted. Any differences across the four census regions will also be highlighted.

[8] The same was true when disaggregating data for Black/African American and Hispanic/Latino educators except for CKT 8 where the average was 4.19 for Hispanic/Latino educators. Disaggregating data by census regions showed importance judgments above 4.2 for all CKT areas and MPs except CKT areas 1 (4.09) and 8 (4.16) for the Midwest and CKT 12 (4.00) for the West.

[9] The same was true when disaggregating data for Black/African American and Hispanic/Latino educators except for CKT 2 which was not in the top four for Black/African American educators (CKT 7 was in the top four). Disaggregating data by census regions produced the same top four CKT areas across regions except for CKT 14 for the West and the same top three MPs across regions except MP 5 for the Midwest.