Instruments to Measure Elementary Preservice Teachers' Conceptions: An Application of the Rasch Rating Scale Model

Cindy Jong
University of Kentucky

Thomas E. Hodges
University of South Carolina

Kenneth D. Royal
North Carolina State University

Rachael M. Welder
Western Washington University

Abstract

This article reports on the development of the Mathematics Experiences and Conceptions Surveys (MECS), a pair of comprehensive instruments designed to measure elementary preservice teachers' dispositions, attitudes, and beliefs about mathematics teaching and learning. In addition to conceptions, MECS aim to capture elementary preservice teachers' related mathematical experiences at various benchmark stages in teacher education programs. The Rasch Rating Scale Model is used to examine the psychometric properties of MECS instruments and to establish six scales that are capable of producing reliable and valid measures. Finally, we present an illustration of how MECS might be used within elementary mathematics teacher preparation.

Introduction

Decades of mathematics education research suggests the strong role conceptions play in influencing the instructional practices teachers use within the classroom (Ball & Cohen, 1999; Ernest, 1989; Richardson, 1996; Wilkins, 2008), and students' opportunities to engage in significant mathematical thinking (Fennema et al., 1996; Staub & Stern, 2002). Researchers and teacher educators often look for opportunities to bring into focus conceptions of mathematics within methods courses and other important experiences within the continuum of teacher education (e.g.

Charalambous, Panaoura, & Philippou, 2009; Quinn, 1997). Despite varying definitions in the literature, we follow Thompson's (1992, p. 130) umbrella description of conceptions as "a general notion or mental structure encompassing beliefs, meanings, concepts, propositions, rules mental images, and preferences." In particular, we focus attention of three specific sub-constructs: attitudes toward mathematics, beliefs about mathematics, and dispositions toward reform practices in mathematics education. In this way of thinking, more positive conceptions of mathematics entail a combination of more positive attitudes, beliefs, and dispositions. Mathematics education literature (e.g. Philipp, 2007) also suggests that while these sub-constructs are related, they not synonymous. In fact, growth in certain sub-constructs may lead to growth in others.

Consensus exists on the critical importance of aligning attitudes, beliefs, and dispositions, with effective mathematics teaching practices within the multidimensional preparation of preservice teachers (Metzger & Wu, 2008). Consequently, the work described here provides an important step towards understanding the development of conceptions within teacher education programs. The importance of such work is two-fold: (a) to quantitatively measure outcomes on conceptions over time, which helps shape learning opportunities for preservice teachers, and (b) to illustrate how the combined efforts of teacher education researchers and psychometricians can produce meaningful measures of critical constructs in teacher education that are missing from teacher quality narratives.

Conceptions for Teaching Mathematics

Broadly, our work is situated within the conceptual framework for assessment teacher education developed by Cochran-Smith and Boston College Evidence Team (2009), which was later adapted to describe the interwoven relationship between aspects of teacher learning and related teacher education experiences. This framework highlights the complex, non-linear process for learning to teach, beginning as K-12 students, continuing into the teaching profession, and influenced by classroom, school, community, and university contexts and experiences. Research in mathematics education suggests that efforts to improve teacher quality should focus on transformative learning experiences that influence teaching practices within varying contexts (Ball & Cohen, 1999). In part, this work focuses on conceptions of the nature of mathematics and how one teaches and learns mathematics (Ball, Lubienski, & Mewborn, 2001). From this perspective, instructional practice serves as a mediator between teacher characteristics (such as conceptions) and student learning (Fennema & Franke, 1992). Consequently, understanding the conceptions teachers hold becomes an important element in understanding the learning opportunities that exist for students. As Wilkins (2008) demonstrated, there is a strong link between beliefs, attitudes, and the extent to which teachers employ reform practices in their classrooms.

Based upon prior work in mathematics education (Philipp et al., 2007; Thompson, 1992; Wilkins, 2008), we operationalize conceptions to include dispositions, beliefs, and attitudes, as each plays a key role in understanding the instructional practices that are adopted by preservice teachers. We define each of the terms in the following way:

Beliefs

"Psychologically held understandings … about the world" (Philipp, 2007, p. 259). Beliefs tend to be true/false oriented and context dependent. Further, beliefs are more cognitive, felt less intensely, and harder to change than attitudes (Thompson, 1992). Research indicates the important role beliefs play in the opportunities students have to engage in significant mathematical thinking (e.g. Fennema et al., 1996; Staub & Stern, 2002), and the integration of particular instructional materials and strategies (e.g. Walen, Williams, & Garner, 2003). The beliefs sub-scale focuses on beliefs about what mathematics is, what it entails, and its usefulness in out-of-school contexts.

Attitudes

Judgments made about particular places, events, people, or objects. Attitudes are, to some degree, either positive or negative (Breckler & Wiggins, 1992). Attitudes change more quickly and are less cognitive than beliefs (Philipp, 2007). Consequently, although beliefs can change within the context of a mathematics methods course (Philipp et al., 2007), we might anticipate shifts in attitudes within shorter timeframes that either beliefs or dispositions. The attitudes sub-scale focuses on participants' enjoyment in engaging in mathematics problem-solving.

Dispositions

A tendency to act in a specified way, to take on a particular position (Bourdieu, 1986; Bourdieu & Nice, 1984). In the case of mathematics teaching and learning, our interests center on our desire to understand how preservice teachers position themselves and their K-12 learning experiences with respect to reform recommendations in elementary mathematics education. We see this as related, but distinct from beliefs, as PSTs may position themselves in ways that contrast professed beliefs. While the beliefs and attitudes sub-scales focus on mathematics content, the dispositions sub-scale maintains a focus on pedagogical aspects of development

Taken together, the three sub-constructs provide an image of participants' conceptions of mathematics teaching and learning, which can be observed throughout teacher preparation to understand the relationship between the evolution of conceptions and experiences known to influence those conceptions. Over the past three decades, researchers have examined teachers' conceptions of mathematics and mathematics teaching and learning (Beswick, 2005; Kajander, 2010; McLeod, 1994; Ross, McDougall, & Hogaboam-Gray, 2003). While these studies have implemented a variety of surveys to capture teachers' conceptions, very few surveys have been consistently used in mathematics education research (Chamberlin, 2010). Much of the prior survey work in mathematics education has focused on beliefs (Perry, Wong, & Howard, 2006; Philipp et al., 2007; Szydlik, Szydlik, & Benson, 2003) and attitudes (e.g. Fennema & Sherman, 1976; White, Way, Perry, & Southwell, 2006) independently, with little, if any, specific attention to dispositions. With the exception of very few studies (e.g., Philipp et al., 2007), recent attention to conceptions (in particular, beliefs; see for example, Grootenboer, 2008) has focused on creating rich descriptions of a small number of participants which are helpful in theory building, but limited in the ability to test theory across a large number of participants in a wide variety of teacher education contexts. Therefore, there is a need for large scale studies that model the relationship between meaningful constructs in teacher development (Adler, Ball, Krainer, Lin, & Novotna, 2005) along with "more robust and rigorous scales" (Grootenboer, Lomas, & Ingram, 2008, p. 4).

To address the need for more comprehensive instrumentation assessing conceptions within the teacher education context, this article reports on the Mathematics Experiences and Conceptions Surveys (MECS) designed to understand the development of elementary preservice teachers' (PSTs) attitudes about mathematics, beliefs about mathematics, and dispositions toward reform mathematics teaching and learning. In addition to conceptions, MECS aims to capture related mathematics education experiences of elementary preservice teachers at various stages in teacher education programs, such as field experiences and mathematics methods courses, to indicate whether such classroom and course-based experiences were aligned with reform recommendations in mathematics. Further, MECS include ample demographic information, which, when combined with measures of conceptions, might be used to leverage learning opportunities within mathematics education coursework. We report on the psychometric properties of the conceptions sub-scales within MECS to evaluate their quality and functioning. Further, we provide examples of MECS outcomes to illustrate MECS' potential to observe conceptions within mathematics teacher preparation.

Methods

Participants and Contexts

The data presented here were obtained from 140 elementary PSTs at three universities in the Eastern United States. Participants were enrolled in an elementary mathematics methods course as part of teacher education programs for initial licensure during the spring 2012 semester. All participants completed MECS-M1 and MECS-M2, during the semester in which they completed their mathematics methods coursework. MECS-M1, designed to be taken at the beginning of mathematics methods coursework, was administered during the first week of each methods course; MECS-M2, designed to be taken at the end of mathematics methods coursework, was administered during the final week of each of the same courses.

Two of the three sampled institutions offer traditional, undergraduate, certification programs in elementary education, which require one to two semesters of mathematics methods courses. PSTs at these universities complete their methods coursework towards the culmination of their programs, generally the semester prior to a full-time student teaching or internship experience. The co-requisite field experiences for these methods courses place PSTs in the classroom for 15-25 full days, where they teach a minimum of ten lessons. Contrastingly, participants at the third institution were graduate students enrolled in a non-traditional, two-year, initial elementary certification program. Here, PSTs complete one mathematics methods course and a co-requisite field experience that requires them to observe ten hours at an elementary school. All participants completed the aforementioned co-requisite field experiences during the semester they completed MECS.

Instrumentation

MECS-M1 and MECS-M2, are the focus of this article. They were designed as pre- and post-surveys to be taken at the beginning and conclusion of mathematics methods coursework. To develop these surveys, we reviewed a variety of existing instruments that measure aspects of beliefs, attitudes, and/or dispositions toward mathematics. In particular, we drew from the Teachers for a New Era surveys (Cochran-Smith & the Boston College Evidence Team, 2009), the Nature and Implications of Mathematics and Science Survey (Salish Final Report, 1997), and the Attitudes Toward Mathematics Inventory (Tapia & Marsh, 2004). We also examined items from the Fennema-Sherman Mathematics Scales (Fennema & Sherman, 1976), since this survey has been the most widely used for decades. We recognize that the Mathematics Anxiety Rating Scale (MARS) has also been widely used in the field of mathematics education (Richardson & Suinn, 1972; Suinn & Winston, 2003) Swars, Daene, & Giesen, 2006; Tooke & Lindstrom, 1998). However, since mathematics anxiety is not an area of focus in our work, MARS items were not considered in regards to MECS development.

What resulted from our review was a set of items aimed at understanding PSTs' attitudes toward mathematics, dispositions toward mathematics, beliefs about mathematics, among additional context-based items. In particular, the items we constructed ask PSTs to characterize aspects of mathematics instruction (e.g. use of manipulatives, problem solving, technology, etc.) and report their perceived experiences in mathematics throughout their past K-12 education. In addition to providing candidate characteristics data, these complementary surveys will be able to provide valuable programmatic feedback.

MECS-M1 is designed to measure constructs related to preservice teachers' past K-12 experiences in mathematics, entering beliefs about mathematics, dispositions toward teaching mathematics, and attitudes toward mathematics. MECS-M2 is designed to measure constructs related to preservice teachers' fieldwork experiences, experiences in the mathematics methods coursework, beliefs about mathematics, dispositions toward teaching mathematics, and attitudes toward mathematics. The two instruments consist primarily of five-point Likert-scale items (1 = strongly disagree, 2 = disagree, 3 = uncertain, 4 = agree, and 5 = strongly agree) and four open-ended questions. MECS-M1 includes 65 Likert-scale items and MECS-M2 includes 80 Likert-scale items. Conceptions subscales in both MECS-M1 and MECS-M2 were created with identical constructs (i.e. attitudes, beliefs, and dispositions) containing 14 items. The identical constructs also allowed us to avoid a form of single-method bias and to measure growth after a mathematics methods course and over time. There are also identical items within the social justice and confidence constructs, not discussed in this paper. The remaining items within each instruments consist of various experience subscales. MECS-M1 contains a "past K-12 experiences" subscale, whereas MECS-M2 contains a "field experience" and "math methods" subscale; this explains the difference in the number of items. Table 1 displays sample items across MECS-M1 and MECS-M2. Because there are two versions of the MECS, two separate analyses were performed to evaluate the psychometric properties of the instruments.

Data Analysis

Initial analyses addressed issues of internal validity with item revisions and content validity by having an expert review panel of mathematicians and mathematics educators scrutinize adequacy of items and subscales. The review took place during two roundtable conference sessions and through solicitation made directly by authors of this manuscript.

Rasch Measurement

For a more thorough evaluation of psychometric properties of the MECS the Rasch Rating Scale Model (RRSM, Andrich, 1978) was utilized. According to Ludlow, Enterline, and Cochran-Smith (2008), "Rasch models are used as confirmatory tests of the extent to which scales have been successfully developed according to explicit a priori measurement criteria" (p. 196). Rasch models are preferred by many researchers because they not only address traditional psychometric criteria, but also provide detailed diagnostic information about the quality and structure of the data, but are also capable of investigating person and item interactions separately (Andrich, 2011). Royal (2010) lists six significant limitations of traditional survey analyses, including 1) ordinal rating scales are erroneously treated as interval measures; 2) all items are considered to be equally important; 3) error is assumed to be equal across all measures; 4) data are sample-dependent; 5) traditional methods typically require normally distributed data; and 6) missing data for some items often results in invalid responses for an entire data record. Rasch models overcome each of the aforementioned limitations

Table 1. Sample MECS Items

Survey	Sample Likert-Scale Item	Construct
MECS-M1 MECS-M2	Mathematics includes questioning and explaining. Mathematics is useless in society. Using mathematics is essential to every day life.	Beliefs about Mathematics
MECS-M1 MECS-M2	I enjoy solving mathematics problems. I think mathematics is boring. I look forward to teaching mathematics.	Attitudes toward Mathematics
MECS-M1 MECS-M2	I plan to use hands-on materials (e.g., blocks, cubes, spinners) to teach mathematics. I plan to encourage students to solve mathematical problems in more than one way.	Dispositions toward Teaching Mathematics
MECS-M1	I struggled with mathematics as a K-5 student. My experiences learning mathematics were generally positive as a K-5 student.	Experience as K-12 Learner of Mathematics
MECS-M2	My math methods course(s) emphasized using and managing manipulatives to teach mathematical concepts. My math methods course(s) emphasized connecting mathematics to students' lives outside the classroom.	Mathematics Methods Course Experience
MECS-M2	My cooperating teacher(s) contributed greatly to my knowledge about the teaching and learning of mathematics. Generally, I had positive field-based experiences (practicum, internship, block, etc.) this semester.	Fieldwork Experience

(Royal, 2010). For a thorough overview of Rasch models and the advantages of these models, readers are encouraged to read Bond & Fox (2007) or Engelhard (2013).
encouraged to read Bond & Fox (2007) or Engelhard (2013).

The Rasch Rating Scale Model (Andrich, 1978) was selected as the measurement model, as the RRSM is well-suited for polytomous data that contain the same number of possible response options. According to the RRSM model, the probability of a person n responding in category x to item i, is given by:

$$P_{xni} = \frac{\exp \sum_{j=0}^{x} [\beta_n - (\delta_i + \tau_j)]}{\sum_{k=0}^{m} \exp \sum_{j=0}^{k} [\beta_n - (\delta_i + \tau_j)]} \qquad x = 0, 1, \ldots, m$$

where $\tau_0 = 0$ so that exp β_n is the person's position on the variable, δ_i is the scale value (difficulty to endorse) estimated for each item i and $\tau_1, \tau_2, \ldots, \tau_m$ are the m response thresholds estimated for the m + 1 rating categories. Winsteps (Linacre, 2012) measurement software was used to analyze the data. Winsteps estimated parameters for the model using joint maximum likelihood estimation (JMLE) procedures (Wright & Masters, 1982).

Results

Overall Results of MECS' Conceptions

Psychometric properties of the MECS were evaluated by way of a collective Rasch analysis of the shared conceptions (attitudes, beliefs, and dispositions) items, and by a Rasch analysis of subscales. A principal components analysis (PCA) of standardized residual correlations was performed to assess dimensionality on the entire instrument. Results indicated the primary dimension explained 60.9% of the variance in the data. The largest secondary dimension explained 11.7% of the variation, indicating a hint of a secondary dimension. The Eigenvalue of the secondary dimension was 4.2 indicating a strength of about 4 items. Fit statistics indicated good fit with Mean Square Infit values of .99 for persons and 1.01 for items. Mean Square Outfit values were 1.04 for persons and 1.04 for items. Provided the data fit the RRSM quite well and did not require the deletion of any misfitting persons or items, and given the secondary dimension was proportionally small relative to the primary dimension, evidence was available to suggest the data were sufficiently unidimensional for a Rasch analysis. In other words, the conceptions items found in MECS generally form a single underlying pattern.

Reliability estimates were .87 for persons and .99 for items. These values indicate highly reproducible measures. Separation refers to the number of statistically distinguishable levels that an instrument can identify in a sample (Wright, 1996). The separation index for persons was 2.58, indicating nearly 3 statistically distinguishable levels were discernible within the person data. Item separation values were 9.61 indicating about 10 statistically distinguishable levels were discernible among the items.

With regard to item measures, all items appeared to be functioning well as determined by fit statistics. Measures ranged in values from -1.94 to 2.24, indicating a good spread of logits. Mean Square fit statistics indicated two items fit the criteria for being potentially "noisy", as their values exceeded the .6 to 1.4 suggested ranges (Wright & Linacre, 1994). These two item values ranged from 1.44 to 1.62. Only items that exceed values of 2.0 have the potential to distort the measurement system, so the effects of including these items in the MECS are negligible. Local dependency was investigated by examining the residual item correlations. Values greater than 0.3 are considered potentially dependent (Smith, 2000). Only four items possessed standardized residual correlations exceeding the 0.3 threshold, and these potentially dependent items were correlated with several items. Future revisions of the MECS will seek to revise these four items to better ensure independence. However, as Linacre (2009) points out, "it is not the extent or the size of violation that is crucial, but rather its impact on critical aspects of the measuring system" (p. 12). Because the fit statistics for the items demonstrating potential dependence were within acceptable ranges, we believe the practical impact of this potential dependence is minimal.

Rating scale functioning was also evaluated. Fit indices indicated good fit for each category of the Likert-type scale and stepwise calibrations advanced in an ordinal manner, with exception to the neutral category. While the neutral category of "uncertain" did not negatively affect the quality of measures, future administrations of the MECS will exclude this category for additional psychometric testing and include a forced-choice six-point Likert-scale. While conceptions items were found to be psychometrically unidimensional, our framework recognizes that conceptions are comprised of multiple psychological processes – namely, attitudes, beliefs, and dispositions.

Conceptions Subscales

Relevant psychometric results for the three common subscales on MECS-M1 and MECS-M2 are presented below (Tables 2 and 3). Results indicate the data fit the RRSM well. Reliability and separation measures are quite high for each subscale, except Dispositions toward Teaching Mathematics. Upon further investigation of the item calibrations for the

Table 2. Person Fit Statistics and Reliability Measures

Subscale	Infit Mean Square	Outfit Mean Square	Reliability	Separation
Attitudes toward Mathematics	1.04	1.05	.90	2.98
Beliefs about Mathematics	0.95	1.01	.78	1.90
Dispositions toward Teaching Mathematics	0.98	0.96	.42	0.85

Table 3. Item Fit Statistics and Reliability Measures

Subscale	Infit Mean Square	Outfit Mean Square	Reliability	Separation
Attitudes toward Mathematics	1.00	1.05	.98	6.59
Beliefs about Mathematics	1.00	1.04	.98	6.84
Dispositions toward Teaching Mathematics	1.00	0.96	.87	2.57

Dispositions subscale, it is apparent that the lack of variability is primarily due to only having three items with two of the three being fairly homogenous in terms of their locations on the measurement continuum. To improve reliability estimation and separation, additional, more difficult to endorse items will be added.

Experience Subscales

The MECS-M1 contains a subscale for "past K-12 experiences". The six items that make up this subscale were Rasch analyzed. Results indicate very good data to model fit with Mean Square Infit values of 1.00 for persons and .98 for items, and Mean Square Outfit values of 1.02 for persons and 1.02 for items. Measures were also highly reproducible with reliability estimates of .84 for persons and .97 for items. Point-measure correlations ranged from .64 to .77 for each of the items, and all items were well within the recommended ranges for item fit values.

The MECS-M2 contains subscales for "field experiences" and "math methods course". First, relevant psychometric results are presented for the three items that comprise the "field experiences" subscale. Results indicate marginally good data to model fit with Mean Square Infit values of .91 for persons and .99 for items, and Mean Square Outfit values of .93 for persons and .93 for items. Measures were also moderately reproducible with reliability estimates of .76 for persons and .95 for items. Point-measure correlations ranged from .74 to .85 for each of the items, and all items were well within the recommended ranges for item fit values. While the results for this subscale are psychometric sound, we will continue to improve the quality of the measures with the addition of items.

Next, relevant psychometric results are presented for the 17 items that comprise the "math methods" subscale. Data fit the RRSM very well, with Mean Square Infit values of 1.00 for persons and 1.01 for items, and Mean Square Outfit values of .97 for persons and .97 for items. Reliability estimates were very high, with values of .95 for persons and .98 for items. Point-measure correlations for each item approximated .72, with the exception of one potentially noisy item. This item indicated a hint of off-variable noise (Infit Mean Square value of 1.49), enough to generate a flag, but not enough to merit a revision or removal of the item. Collectively, the "math methods" subscale functions exceptionally well.

Discerning the Evidence for Construct Validity

Measurement scholar Samuel Messick (1989) presented a uniformed conceptualization of construct validity. Messick contends construct validity requires an accumulation and integration of evidence that affects the interpretation or meaning of one's results. Messick's framework identified six aspects of construct validity: content, substantive, structural, generalizeability, external and consequential. Thus, we evaluate the construct validity of the MECS from this perspective.

First, content validity was evaluated by way of item fit statistics. Virtually every item on both versions of the MECS fell within the recommended range of .6 to 1.4. Additionally, because the sample size consisted of 140 preservice teachers, standard errors for each item measure were quite small. Naturally, smaller standard errors lead to greater stability, thus providing some assurance that the construct is stable. Collectively, there is sufficient evidence to support the content aspect of validity.

Second, the substantive aspect of validity was evaluated by way of a principal components analysis (PCA) of standardized residual correlations. Results indicated that 60.9% of the variance was explained by the measures, thus indicating a highly unidimensional construct. Further, data tend to fit the Rasch Rating Scale Model quite well, even without the exclusion of persons who misfit the model's expectations. Collectively, these results present evidence to support the substantive aspect of validity.

Third, structural validity was evaluated by way of investigating respondents' use of the rating scale. Generally, structural calibrations advanced in a stepwise, ordinal manner with the exception of the neutral category. Although the effects of the neutral category did not appear to distort the measures or cause any problems, future administrations of the MECS will not include the neutral category of "uncertain". The fact remains that respondents were able to adequately discern the hierarchical nature of the rating scale and the categories' distinctions appeared to be appropriate for the nature of the questions. Evidence of good rating scale quality was available, which speaks to both the structural and communicative (Lopez, 1996) aspects of validity.

Fourth, the generalizability aspect of validity was evaluated. Reliability estimates were moderate to high for both persons and items for every subscale but one. The "dispositions to teaching mathematics" subscale provided less than desirable reliability estimates via the reliability calculation procedures incorporated in a Rasch analysis. Items will be added in the updated dispositions subscale to include more difficult to endorse. This has the potential to improve reliability estimates due to the additional variability produced by the more distant item location on the empirical construct continuum. Despite the small issue with one subscale, a plethora of evidence was available to support the generalizability aspect of validity.

Finally, we present no evidence of external or consequential validity in this study. Additionally, we did not investigate systematic validity, or the extent to which the construct hierarchy remains stable across various person demographic subgroups. Future research on larger, and more diverse samples will utilize Differential Item Functioning (DIF) techniques in addition to the traditional Rasch analysis to further discern systematic validity. In conclusion, the collective sum of empirical evidence suggests the MECS is an instrument capable of producing quality measures when targeted to a similar sample frame.

Discussion

In the following sections, we discuss the results of the psychometric analyses of MECS instrumentation. Further, we provide an illustrative example of how MECS have been used to address significant questions in the preparation of elementary mathematics teachers.

Psychometric Analyses

Establishing psychometrically sound instruments is an ongoing, iterative process. As MECS instruments are adopted at additional institutions and we expand our data set, we will continue to analyze and refine the items and constructs of our MECS instruments. Initial analyses of Likert-scale items have identified six highly reliable sub-scales within our instruments: Attitudes toward Mathematics, Beliefs about Mathematics, Dispositions toward Teaching Mathematics, K-12 Learner of Mathematics, Mathematics Methods Course Experiences, and Field Experience. These results highlight the fact that we were able to achieve high reliability on multiple sub-constructs of teachers' conceptions within two complementary surveys. Given the need to establish more robust scales and sophisticated instruments, these findings are significant to understanding how PSTs conceptions are developed with respect to reform recommendations in mathematics. Furthermore, high reliability of measures of Dispositions toward Teaching Mathematics alongside Attitudes toward Mathematics and Beliefs about Mathematics provided promising findings regarding the ability of the

MECS to provide a more comprehensive picture of PSTs' conceptions.

Subsequent MECS survey data will be evaluated to measuring change across two points in time. Using a relatively simple item and rating scale threshold anchoring procedure, data from pre- and post- administrations of the MECS can be equated onto the same exact scale for direct comparison (Wright, 1996, 2003). Traditional statistical procedures will then be utilized to both describe the differences between pre- and post- responses, as well as discern the distribution and magnitude of these changes (if any).

Illustrations of the Potential Uses of MECS

MECS data may be used to address a variety of questions focused on conceptions in relation to mathematical experiences. To provide an illustration, we draw from two administrations of MECS at two institutions in the United States, which included a total of 77 cases. A multitude of studies in mathematics education indicate the influence primary and secondary mathematics experiences have on PSTs' attitudes toward and beliefs about mathematics (C. A. Brown & Borko, 1992; S. I. Brown, Cooney, & Jones, 1990; Raymond, 1997). Consequently, we sought to determine if MECS could capture the relationship between PSTs' characterizations of their K-12 mathematics experiences and their entering attitudes about mathematics. Further, we sought to determine if changes in attitudes occurred over the duration of the methods course and whether those changes could be explained by certain types of methods course experiences.

A Pearson's r was computed using the logit values produced from the RRSM, with the variable K12_exp having a strong positive correlation with the variable attitudes_pre, $r(75) = .605$, $p < .01$. We see this as a promising finding as many existing studies have used qualitative methods to observe such links.

Of particular interest were those PSTs which reported relatively "unfavorable" K-12 experiences, representing a population of PST which may be of increased concern. That is to say, PSTs who enter teacher education programs with negative past K-12 mathematics experiences might need additional support to develop more favorable conceptions. Using a cutoff K12_exp logit score of zero to identify unfavorable experiences, 28 of 77 (36%) cases were selected. A single sample t-test was calculated using the variable attitudes_change, representing the difference in logit values (attitudes_post – attitudes_pre). There was a statistically significant gain ($M = 0.781$, $SD = 0.252$) in attitudes among those with unfavorable K-12 experiences, $t(27) = 3.098$, $p < .01$.

To determine whether the changes in attitudes could be explain, in part, by methods course experiences captured in MECS, we used a backward elimination regression model technique. The model was loaded with mathematics methods pedagogy (MM_ped), materials (MM_mat), environment (MM_env), and students (MM_stu) – constructs used to identify foci in methods course experiences. When attitudes_change was predicted it was found that MM_stu ($\beta = -0.595$, $p = .089$) and MM_ped ($\beta = 1.440$, $p = 0.022$) were the best predictors. The overall model fit was $r2 = .449$. While this regression model was limited, due in part to the small sample size, it identified two potentially influential methods course experiences. In our example, MECS was able to observe: (a) the link between K-12 experiences and entering attitudes; (b) changes in attitudes over the duration of a mathematics methods course among those with unfavorable experiences; and (c) potential links between changes and particular methods course experiences.

The examples here are illustrative, rather than exhaustive, of the types of questions that may be asked of MECS data. A wide array of questions focused on the relationship between conceptions and mathematical experiences may be explored with linearized MECS data. However, we acknowledge the limitations imposed by surveys containing Likert-type items alone and have therefore included open-ended response items on each MECS instrument. We also realize that there are challenges associated with measuring conceptions in mathematics education (Leder & Forgasz, 2006). Although initial analyses have focused on establishing construct validity, our continued analyses will include a qualitative investigation of the responses to open-ended items for triangulation purposes.

Future Research

We believe that the MECS will be useful to researchers across a variety of contexts in studying elementary preservice teachers' attitudes, beliefs, and dispositions toward mathematics. First, unlike other surveys designed to measure teacher attitudes and beliefs, our instrument collects extensive demographic information, such as participants' past experiences as K-12 learners of mathematics. MECS administration to large, diverse sets of teachers across multiple institutions may inform group differences in teachers' conceptions of mathematics.

Furthermore, using MECS affords opportunities to measure changes in teacher conceptions from entry into a teacher education program into the first years of teaching. To our knowledge, MECS are the only instruments designed to observe the development of teachers' conceptions about mathematics over time. MECS is another critical step providing teacher educators with access to a knowledge base to inform the improvement of their programs (Berk & Hiebert, 2009). Baseline beliefs, attitudes, and dispositions of

PSTs can be measured by the MECS-M1 prior to the mathematics methods coursework. Changes in conceptions can then be measured by administering variations of MECS as elementary PSTs meet significant benchmarks throughout their programs, such as the completion of methods courses or student teaching.

Using multiple testing points over time, researchers can to study the impact of particular experiences within and across teacher education programs. This information may help teacher educators understand how elementary PSTs' attitudes, beliefs, and dispositions are formed throughout specific aspects of teacher education and identify factors that contribute to PSTs having certain conceptions that may influence their future teaching practices. Ultimately, we support making the development of teacher conceptions a goal of teacher education programs. This work has the potential to help teacher educators evaluate and improve their programs by identifying and implementing strategies that will positively impact the beliefs, attitudes, and dispositions of their elementary preservice teachers. While mathematics has been the central content area of discussion in this paper, we believe that this work can be used as a roadmap to extend conceptions work across other content areas, providing additional data for teacher educations programs.

References

Andrich, D. (1978). A rating formulation for ordered response categories. *Psychometrika, 43*, 561-573.

Andrich, D. (2011). Rating scales and Rasch measurement. *Expert Review of Pharmacoeconomics & Outcomes Research, 11*(5), 571-585.

Adler, J., Ball, D., Krainer, K., Lin, F. L., & Novotna, J. (2005). Reflections on an emerging field: Researching mathematics teacher education. *Educational Studies in Mathematics, 60*(3), 359-381.

Ball, D. L. (1990). The mathematical understandings that prospective teachers bring to teacher education. *Elementary School Journal, 90*(4), 449-466.

Ball, D. L., & Cohen, D. K. (1999). Developing practice, developing practitioners: Toward a practice-based theory of professional education. In G. Sykes & L. Darling-Hammond (Eds.), *Teaching as the learning profession: Handbook of policy and practice* (pp. 3-32). San Francisco: Jossey Bass.

Ball, D. L., Lubienski, S. T., & Mewborn, D. S. (2001). Research on teaching mathematics: The unsolved problem of teachers' mathematics knowledge. In V. Richardson (Ed.), *Handbook of Research on Teaching* (4 ed., pp. 433-456). New York: Macmillan.

Berk, D., & Hiebert, J. (2009). Improving the mathematics preparation of elementary teachers, one lesson at a time. *Teachers and Teaching: Theory and Practice, 15*(3), 337-356.

Beswick, K. (2005). The beliefs/practice connection in broadly defined contexts. *Mathematics Education Research Journal, 17*(2), 39-68.

Bond, T. G., & Fox, C. M. (2007). *Applying the Rasch Model. Fundamental measurement in the human sciences*, 2nd edition. Lawrence Erlbaum Associate.

Bourdieu, P. (1986). The forms of capital. *Handbook of theory and research for the sociology of education*, 241-258.

Bourdieu, P., & Nice, R. (1984*). Distinction: A social critique of the judgement of taste.* New York: Harvard University Press.

Breckler, S. J., & Wiggins, E. C. (1992). *On defining attitude and attitude theory: Once More with feeling. Attitude structure and function*, 407-427. Florence, KY: Psycholgy Press.

Brown, C. A., & Borko, H. (1992). Becoming a mathematics teacher. In D. Grouws (Ed.), *Handbook of research on* *mathematics teaching and learning* (pp. 209 - 239). New York: Macmillan.

Brown, S. I., Cooney, T. J., & Jones, D. (1990). Mathematics teacher education. *Handbook of research on teacher education*, 639-656.

Chamberlin, S. A. (2010). A review of Instruments Created to Assess Affect in Mathematics. *Journal of Mathematics Education, 3(*1), 167-182.

Cobb, P., Confrey, J., diSessa, A., Lehrer, R., & Schauble, L. (2003). Design Experiments in Educational Research. *Educational Researcher, 32(*1), 9-13.

Cochran-Smith, M., & the Boston College Evidence Team. (2009). Re-Culturing teacher education: Inquiry, evidence, and action. *Journal of Teacher Education, 60*(5), 458-468. doi: 10.1177/0022487109347206

Engelhard, Jr., G. (2013). Invariant measurement: Rasch measurement in the social, behavioral and health sciences. Routledge.

Fennema, E., Carpenter, T. P., Franke, M. L., Levi, L., Jacobs, V. R., & Empson, S. B. (1996). A longitudinal study of learning to use Children's thinking in mathematics instruction. *Journal for Research in Mathematics Education, 27*(4), 403-434.

Fennema, E., & Franke, M. L. (1992). Teachers' knowledge and its impact. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 147–164). New York: Macmillan.

Fennema, E., & Sherman, J. A. (1976). Fennema-Sherman Mathematics attitudes scales: Instruments designed to measure attitudes toward the learning of mathematics by females and males. *Journal for Research in Mathematics Education, 7*(5), 324-326.

Grootenboer, P. (2008). Mathematical belief change in prospective primary teachers. *Journal of Mathematics Teacher Education, 11*(6), 479-497.

Henry, G. T., Purtell, K. M., Bastian, K. C., Fortner, C. K., Thompson, C. L., Campbell, S. L., & Patterson, K. M. (2013). The effects of teacher entry portals on student achievement. *Journal of Teacher Education.*

Kajander, A. (2010). Elementary mathematics teacher preparation in an era of reform: The development and assessment of mathematics for teaching. *Canadian Journal of Education, 33*(1), 228-255.

Linacre, J. M. (2012). WINSTEPS® (Version 3.70). Beaverton, OR: winsteps.com.

Linacre, J. M. (2009). Local independence and residual covariance: A study of olympic figure skating ratings. *Journal of Applied Measurement, 10*(2), 1-13.

Ludlow, L. H., Enterline, S. E., & Cochran-Smith, M. (2008). Learning to teach for social justice-beliefs scale: An application of Rasch measurement principles. *Measurement and Evaluation in Counseling and Development, 40(*4), 194.

Marbach-Ad, G., & McGinnis, J. R. (2009). Beginning mathematics teachers' beliefs of subject matter and instructional actions documented over time. *School Science and Mathematics, 109*(6), 338-354.

McLeod, D. B. (1994). Research on affect and mathematics learning in the JRME: 1970 to the present. *Journal for Research in Mathematics Education, 25(*6), 637-647.

Messick, S. (1989). Meaning and values in test validation: The science and ethics of assessment. *Educational Researcher, 18*(2), 5-11.

Metzger, S. A., & Wu, M. J. (2008). Commercial teacher selection instruments: The validity of selecting teachers through beliefs, attitudes, and values. *Review of Educational Research, 78*(4), 921.

Pajares, M. F. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct. *Review of Educational Research, 62*(3), 307-332.

Perry, B., Wong, N. Y., & Howard, P. (2006). Comparing primary and secondary mathematics teachers' beliefs about mathematics: Mathematics learning and mathematics teaching in Hong Kong and Australia. *Mathematics Education in Different Cultural Traditions- A Comparative Study of East Asia and the West,* 432-448.

Philipp, R. A. (2007). Mathematics teachers' beliefs and affect. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (Vol. 1, pp. 257-315). Charlotte, NC: Information Age Publishing.

Philipp, R. A., Ambrose, R., Lamb, L. L. C., Sowder, J. T., Schappelle, B. P., Sowder, L., . . . Chauvot, J. (2007). Effects of early field experiences on the mathematical content knowledge and beliefs of prospective elementary school teachers: An experimental study. *Journal for Research in Mathematics Education, 38*(5), 438-476.

Raymond, A. M. (1997). Inconsistencies between a beginning elementary school teacher's mathematics beliefs and teaching practice. *Journal for Research in Mathematics Education, 28*(5), 550 - 576.

Richardson, F. C., & Suinn, R. M. (1972). The mathematics anxiety rating scale: Psychometric data. *Journal of Counseling Psychology, 19*(6), 551-554.

Ross, J. A., McDougall, D., & Hogaboam-Gray, A. (2003). A survey measuring elementary teachers' implementation of standards-based mathematics teaching. *Journal for Research in Mathematics Education, 34*(4), 344-363.

Royal, K. D. (2010). Making meaningful measurements in survey research: A demonstration of the utility of the Rasch Model. *IR Applications, 28*(1), 1-16.

Salish Final Report. (1997). Secondary science and mathematics teacher preparation programs: Influences on new teachers and their students (pp.

217). Washington, D.C.: Office of Educational Research and Improvement.

Sfard, A. (1991). On the dual nature of mathematical conceptions: Reflections on processes and objects as different sides of the same coin. *Educational Studies in Mathematics, 22*(1), 1-36.

Smith, R. M. (2000). Fit analysis in latent trait measurement models. *Journal of Applied Measurement, 1*(2), 199-218.

Staub, F. C., & Stern, E. (2002). The nature of teachers' ppedagogical content beliefs matters for students' achievement gains: Quasi-Experimental evidence from elementary mathematics. *Journal of Educational Psychology, 94*(2), 344-355. doi: 10.1037//0022-0663.94.2.344

Suinn, R. M., & Winston, E. H. (2003). The mathematics anxiety rating scale, a brief version: Psychometric data. *Psychological Reports, 92*(1), 167-173.

Szydlik, J. E., Szydlik, S. D., & Benson, S. R. (2003). Exploring changes in pre-service elementary teachers' mathematical beliefs. *Journal of Mathematics Teacher Education, 6* (3), 253-279. doi: 10.1023 / A: 1025155328511

Tapia, M., & Marsh, G. E., II. (2004). An Instrument to Measure Mathematics Attitudes. *Academic Exchange Quarterly, 8*(2).

Thompson, A. G. (1992). Teachers' beliefs and conceptions: A synthesis of the research. In D. A. Grows (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 127 - 146). New York: Macmillan.

Walen, S. B., Williams, S. R., & Garner, B. E. (2003). Pre-service teachers learning mathematics using calculators: A failure to connect current and future practice. *Teaching and Teacher Education, 19*(4), 445-462.

White, A. L., Way, J., Perry, B., & Southwell, B. (2006). Mathematical attitudes, beliefs and achievement in primary pre-service mathematics teacher education. *Mathematics Teacher Education and Development, 7*, 33-52.

Wilkins, J. L. M. (2008). The relationship among elementary teachers' content knowledge, attitudes, beliefs, and practices. *Journal of Mathematics Teacher Education, 11*(2), 139-164.

Wright, B. D. (1996). Reliability and separation. *Rasch Measurement Transactions, 9*(4), 472.

Wright, B. D., & Linacre, J. M. (1994). Reasonable mean-square fit values. *Rasch Measurement Transactions, 8*, 370.

Wright, B. D., & Masters, G. N. (1982*). Rating scale analysis*. Chicago, IL: Mesa Press.