DEVELOPING THE MATHEMATICS EDUCATION OF ENGLISH LEARNERS SCALE (MEELS)

Anthony Fernandes
University of North Carolina – Charlotte
anthony.fernandes@uncc.edu

Laura McLeman
University of Michigan – Flint
lauramcl@umflint.edu

In this paper, we describe the initial stage of reliability and validity testing for the Mathematics Education of English Learners Scale (MEELS), which is designed to measure preservice teachers' beliefs about the mathematics education of English learners. To address the content validity, we consulted with experts within the field of mathematics education to assess the relevance and representation of specific items. We used Principal Component Factor Analysis to determine construct validity. Finally, we tested reliability of the factors using Cronbach's coefficient alpha. After sharing the findings from these analyses, we describe the next stages in the development of MEELS.

Keywords: Beliefs; Measurement; Equity and Diversity; Teacher Education–Preservice

The growing population of English Learners (ELs) in schools across the country has made it necessary to prepare all teachers, including mathematics teachers, to work with EL students in their classrooms (Costa, McPhail, Smith, & Brisk, 2005; Lucas & Grinberg, 2008). Given that beliefs are "lenses that affect one's view of some aspect of the world or as dispositions toward action" (Philip, 2007, p. 259), it is important for educators to know what preservice teachers (PSTs) believe about the mathematics education of ELs, especially when they assume a deficit perspective. As a first step in challenging these beliefs, the goal of our study is to develop a valid and reliable instrument, called the Mathematics Education of English Learners Scale (MEELS), which will measure the beliefs that PSTs have about issues related to the teaching and learning of mathematics to ELs. In this paper, we will describe the first stage analysis for validity and reliability that was conducted after 334 PSTs responded to MEELS.

Literature Review

A review of the literature revealed that there are a number of surveys in the areas of PSTs' or teachers' beliefs about diversity and multiculturalism, language attitudes, and inclusion of ELs. One such survey is the Cultural Diversity Awareness Inventory (CDAI; Henry, 1986), which consisted of 28 statements. The CDAI was developed to measure general cultural awareness that educators had about young children from culturally diverse backgrounds. The inventory was based on the understanding of "culture" that included five areas: (1) values and beliefs, (2) communication, (3) social relationships, (4) food and diet, and (5) dress. Despite the lack of information regarding the validity and reliability of the CDAI, it has been used by other researchers (e.g., Davis & Turner, 1993; Larke, 1990) to assess the cultural sensitivity of elementary PSTs.

Pohan and Aguilar (2001) designed two measures to elicit the personal and professional beliefs that educators had about diversity. In their survey, they extend traditional definitions of diversity of race and/or ethnicity to include other marginalized groups based on social class, gender, religion and sexual orientation. The researchers conjectured that there were some beliefs in the personal and professional spheres that could be in conflict. For example, an educator might believe that bilingualism was good in the current diverse society; however, they may not approve of public money being spent in maintaining bilingual programs. The reliability and validity of both scales were extensively tested with samples that included PSTs, graduate students, and practicing educators from rural and urban schools.

Byrnes and Kiger (1994) designed the Language Attitudes to Teaching Scale (LATS) that contained 13 items about the beliefs that teachers had about EL students. A factor analysis yielded 3 factors tied to the politics of language, intolerance of EL students and language support. The researchers carried out face and construct validity, and the subscale reliabilities, using Cronbach's alpha, ranged from 0.60 to 0.72.

Van Zoest, L. R., Lo, J.-J., & Kratky, J. L. (Eds.). (2012). *Proceedings of the 34th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education.* Kalamazoo, MI: Western Michigan University.

Discussion about best practices for EL students in the content area have consistently advocated for inclusion of these students in mainstream classes along with the use of Sheltered Instruction practices (Echevarria, Vogt, & Short, 2008). This aspect has motivated researchers to examine the beliefs that teachers have towards the inclusion of ELs in their content classes. For example, Reeves (2006) investigated the beliefs that secondary teachers had about having EL students in their classroom through a self-designed survey. Walker, Shafer, and Iliams (2004) used a survey to assess mainstream teachers attitudes towards EL students and how these attitudes varied across schools where there were few ELs, a rapidly growing EL population, and ones where ELs were predominant. In both studies, only the face validity of the instrument was discussed.

All of the surveys reviewed, only a sample of which were discussed in this section, measured teachers' and/or PSTs' beliefs about cultural diversity, their attitudes towards linguistic diversity, and inclusion of EL students in their classroom. The instruments were broad in scope and encompassed teachers and PSTs in all subject areas. Reliability and validity were done only in two instruments (Byrnes & Kiger, 1994; Pohan & Aguilar, 2001). In all other cases, researchers developed surveys to examine the impact of an intervention, but validity and reliability of the instruments were either not discussed or the researchers only used participants' feedback to determine the clarity of the questions (face validity). Absent from these descriptions were other forms of validity like content, construct and criterion, and measures of reliability such as internal consistency reliability and test-retest reliability. Thus we need an instrument that will capture the different dimensions of the construct, have items that are consistent on each dimension, and that remains stable over time.

Beyond the dearth of statistically demonstrated valid and reliable surveys, we also did not find a survey that measured discipline specific beliefs with respect to the mathematics education of ELs. According to Cooney Shealy, and Arvold (1998) and Philip (2007), beliefs are tied to the context and this could be very different for different situations. In mathematics, as compared to other subjects, it is possible that the PSTs may assume that language plays a minimal role in the teaching and learning of mathematics, despite evidence to the contrary (e.g., Moschkovich, 2010). Teachers who assume that language plays a minimal role in the teaching and learning of mathematics then may be less likely to adjust their teaching in order to accommodate ELs. In fact misconceptions about language lead to a high proportion of ELs being labeled with a learning disability when they can converse in English, but struggle with the content, which involves academic English (Gandára & Contreras, 2009). With these considerations, the goal of our study was to design a valid and reliable measure that examined the beliefs that PSTs had about the mathematics education of ELs.

Theoretical Perspective

The overall framing of MEELS and the items in particular were guided by non-deficit views about ELs and their communities (Civil, 2007; Moschkovich, 2010). According to Moschkovich, "deficit models stem from assumptions about learners and their communities based on race, ethnicity, SES (socio-economic status), and other characteristics assumed to be related in simple, and typically negative ways to cognition and learning in general" (p. 11). Non-deficit models, on the other hand, assume that the EL students are part of different Discourse communities and have valuable resources that are assets which teachers can use to develop the students' knowledge (Civil, 2007; Moschkovich, 2010). For example, mathematical algorithms that students may have studied in other countries could be welcomed by the teacher, viewed as a resource, and shared with other students in the class. Further, the communities and parents of the ELs are also seen to have valuable knowledge that could be utilized in the classroom (Civil, 2007). Thus in framing and later scoring the items we assumed, for example, that bilingualism was an asset rather than a hindrance to EL students and that parents from all communities fundamentally cared about the intellectual development of their children, even if this was not visible to the teacher or did not adhere to a preconception of what that caring should look like (e.g., attending parent-teacher conference, volunteering in the classroom, etc.). Overall, the items for MEELS were drawn from recommendations from research, interactions with other mathematics educators, and other diversity surveys [refer to

McLeman & Fernandes (under review) for a review of the literature that guided the item development of MEELS].

Methods

Sample

MEELS was administered to 334 PSTs, about 75% of whom were located at one university in the southeast of the United States. Of the 330 responses we analyzed (4 were determined to be outliers), about 86.1% were female and about 84.8% self-reported their race as "White, not of Hispanic origin." Furthermore, a little more than 70% of the participants wanted to teach grades K–5, while about the same percentage of participants had no prior teaching experience. Close to 72% had been exposed to issues involving ELs in prior coursework and about 77% had been involved in some type of field experience during their teacher preparation program. Finally, about 92% of the participants self-reported that they were not fluent in a language other than English, though about 85% did report that they had some experience in learning a second language.

Instrument

The first iteration of MEELS consisted of two sections: the first was comprised of 8 demographic items while the second had 26 items related to the teaching and learning of mathematics to ELs. Each item was measured on a 5-point Likert-type scale: Strongly Disagree (1), Disagree (2), Undecided (3), Agree (4), and Strongly Agree (5). At the end of the survey, participants were asked to answer three open-ended questions to ascertain the readability and clarity of the instrument. Content validity was determined by sending an initial set of items to 10 mathematics education experts in the area of ELs. Based on their feedback, we modified some of the items before the first pilot of MEELS. For example, the language in several items was modified and additional items were included.

MEELS was administered online through SurveyShare (http://www.surveyshare.com) and thereafter the data were downloaded to SPSS 17 for analysis. Using our theoretical perspective as a basis, we reverse scored certain items that would reflect deficit beliefs based on our reading of the literature. For example, the item "Some EL's home culture negatively impacts their math learning" was reverse coded with a PST strongly agreeing scoring a 1 and strongly disagreeing scoring a 5. Since beliefs can only be inferred, it was conjectured that a PST who believed that an EL's home culture could negatively impact their mathematical learning would be less open to seeing certain ELs' home culture as resource in the classroom. In total, 14 of the 26 items on the survey were reverse coded; these are shown with an r in Table 1. Note that there are 23 items in Table 1 as three items were dropped in later analysis. One of the dropped items was reverse coded.

Statistical Analyses

Our main goal was to establish construct validity of MEELS through factor analysis and examine reliability of the resulting subscales associated with the factors (also referred to as internal consistency). We first scanned the data for outliers using Mahalanobis distance. Next, we examined the coefficient matrix, which summarizes the interrelations between the set of items. We performed Bartlett's test of sphericity to ensure that the correlation matrix was not an identity matrix, which would indicate that there was no relationship between the items. We examined the Kaiser-Meyer-Olkin (KMO) statistic that indicates if the sample size was adequate relative to the 26 items in the instrument. In addition to the overall KMO, we also examined the anti-image correlation matrix for a Measure of Sampling Adequacy (MSA) for the individual items. Pett, Lackey, and Sullivan (2003) recommended that the individual MSA (numbers along the diagonal of the anti-image correlation matrix) should be greater than 0.60 to ensure the presence of underlying factors.

After this preliminary analysis, we proceeded to identify clusters of inter-correlated items, usually referred to as factors, which would indicate the various dimensions related to our construct of the mathematics education of ELs. It is important to note that in developing the items we had some conjectures

Van Zoest, L. R., Lo, J.-J., & Kratky, J. L. (Eds.). (2012). *Proceedings of the 34th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education.* Kalamazoo, MI: Western Michigan University.

based on the literature about what these dimensions might be. We used exploratory factor analysis as opposed to confirmatory factor analysis because we were uncertain about the dimensions, and in the process also wanted to ensure construct validity. There are different factor extraction methods in SPSS 17, though Principal Component Analysis (PCA) and Principal Axis Factoring (PAF) are most widely used (Pett, Lackey, & Sullivan, 2003). In trying to determine the best method for our purposes, we relied on the advice of Pett, Lackey, and Sullivan to start with a preliminary solution using PCA, refine the solution by examining the items that load on the various factors (where *load* refers to the correlation between an item and factor), and then develop a preliminary solution. This solution is then compared to the PAF solution on the same matrix and the final solution is "one that is the best fit and that makes the most intuitive sense" (Pett, Lackey, & Sullivan, 2003, p. 115). Once we obtained our factors we worked out the internal consistency of the items that made up a particular factor using Cronbach's coefficient alphas.

Results

The Mahalanobis distance for multivariate data ($p < .001$) (Stevens, 1992; Tabachnick & Fidell, 1996) revealed four outliers that were dropped from the subsequent analysis. Thus the total number of responses examined was 330. Bartlett's test of sphericity was significant (chi-square = 1951.35, df = 253, $p < 0.001$) indicating that there were relationships between the items. The Kaiser-Meyer-Olkin of 0.854 was greater than 0.7 suggesting a sufficient sample size (Kaiser, 1974). The diagonal of the anti-image correlation matrix yielded the Measures of Sampling Adequacy (MSA) which ranged from 0.54 to 0.91, with most values greater than 0.7 and the off-diagonal absolute values were small, thus suggesting that the matrix was factorable (Pett, Lackey, & Sullivan, 2003).

Factor Analysis

Initial PCA suggested 7 factors that satisfied Kaiser's rule (Kaiser, 1974) with eigenvalues greater than 1. Note that our goal in factor analysis was to reduce the number of variables (items in the instrument) into a smaller number of factors that would account for as much of the variation between the individual variables. Towards this end, we chose factors with the largest eigenvalues that would explain more of the variance than an individual item. In exploratory factor analysis, once the initial solution is obtained, variables generally tend to load highly on a factor and have small loadings on others. By rotating the factor axes, the loading of a variable is increased on one of the extracted factors and is minimized on the other factors (also called a simple structure). Thus rotation increases the interpretability of the factor as a group of items load highly on it. In trying to determine the number of factors to retain after rotation, Pett, Lackey, and Sullivan (2003) suggested retaining the fewest number of factors that explained at least 50% of the variance and that the factors make intuitive sense in the given context. With this in mind, we sought to have at least 3 items load on a factor with a loading greater than 0.3, all the factors account for at least 50% of the variance, and that the resulting item-factor correlation matrix achieve simple structure. A closer examination of the rotated matrix with loadings that were more than 0.30, suggested factor 6 comprised of items 18, 23, and 33, and items 23 and 29 loaded on factor 7. Items 33 and 29 loaded only on factors 6 and 7, respectively, while items 18 and 23 also loaded on factors 4 and 5. Since items 29 and 33 only loaded on 6 and 7, we decided to rerun the analysis without these items. A subsequent PCA with Varimax rotation yielded 6 factors with eigenvalues more than 1 and several items that loaded on multiple factors. To make it easier for interpretation, we decided to look at items that loaded 0.40 or more on a factor. Still there were 5 items that loaded on multiple factors and after dropping these items one at a time in further analyses, a decision was made to eliminate item 26. A rerun of the PCA with Varimax rotation, in this case, extracted 5 factors with eigenvalues more than 1. The factors accounted for about 52% of the variance in these items. Table 1 displays the 23 items with loadings greater than 0.4 that loaded on 5 factors in an almost simple structure (three items loaded on multiple factors). After we obtained this initial solution, we followed Pett, Lackey, and Sullivan's (2003) suggestion and ran the PAF with Varimax rotation (with items 26, 29, and 33 dropped). This yielded 4 items that did not load (>0.40) on any of the 5 factors; one of the factors had only 2 items, and another factor that was difficult to interpret. Thus we retained the 5 factor solution obtained with PCA and Varimax rotation as it was the one that most aligned with our extraction criteria (modifying the loadings from >0.30 to >0.40). For the three items that loaded on multiple factors, we decided to associate them with the factor for which a higher loading was demonstrated. Note that a consideration of loadings more than 0.45 would have yielded a simple structure for the item-factor loading matrix shown in Table 1. Note that since we approximated simple structure with the simpler orthogonal rotation (Varimax), we avoided using oblique rotation in the above analysis.

Table 1: Factor Loadings

	Items	**T**	**LSC**	**F**	**LM**	**C**
17	Open to teaching ELs math.	0.52				
22	Open to integrating EL's background in math.	0.62				
27	Adjust the language on math problems.	0.65				
28	Focus on the language skills.	0.67				
30	Accept EL's non-verbal communication.	0.70				
31	Accept alternative math algorithms.	0.55				
36	Open to using alternative math assessments.	0.57		(0.41)		
11	*Fluency* in more than one language.		0.65			
12r	English as only language of instruction		0.75			
13r	More important for beginning ELs to learn English.		0.45			
14	Open to use of native language.	(0.44)	0.56			
15r	Native language use hampers learning English.		0.50			
16	State math tests in different languages.		0.50			
18r	Rich math discussions.			0.61		
32r	Teach ELs and non-ELs in the same way.			0.62		
34r	Accommodations are unfair.			0.53		
35r	Same standards for ELs and non-ELs.			0.79		
19r	Math ideal for transition of beginning ELs.				0.71	
20r	Conversational language.			(0.40)	0.64	
21r	Math is not language intensive.				0.77	
23r	EL's home culture.					0.57
24r	Parents.					0.72
25r	Some ethnicities better at math.					0.71

T=Teaching, LSC=Language in school context, F=Fairness, LM=Language and Mathematics, C=Culture.

Based on the items that loaded on the 5 factors, and paying particular attention to the items that displayed higher loadings, we labeled the factors—Beliefs about teaching (T; 7 items), Beliefs about language in the school context (LSC; 6 items), Beliefs about fairness (F; 4 items), Beliefs about the interconnection of language and mathematics (LM; 3 items) and Beliefs about culture (C; 3 items). The alpha coefficients are given in Table 2.

Table 2: Cronbach's Coefficient Alphas

	Alpha
Teaching	.79
Language in School Context	.73
Fairness	.66
Language and Mathematics	.59
Culture	.48

Discussion

The development of a survey and establishing validity and reliability is an evolving process. In this first iteration, we determined content and construct validity along with internal consistency reliability. For the content validity we consulted with experts within the field of mathematics education to assess the relevance and representation of the items we included in our instrument. We used Principal Component Factor Analysis (PCA) along with Varimax rotation to determine the construct validity of the underlying factors that impact PSTs' beliefs about the mathematics education of ELs. To determine reliability of the factor subscales, we used Cronbach's coefficient alpha. These analyses point to the next steps in the refinement of MEELS.

Next Steps

Cronbach's coefficient alphas for the two subscales LM and C are lower than 0.6, which is poor (Burns & Burns, 2008) and will need to be addressed in the next phase of refinement and testing of MEELS. The presence of only 3 items in each of these subscales could be a possible reason for the low alphas. Thus one course of action would be to increase the number of items in this category. However the inclusion of additional items will need to take into the consideration the connection that issues of culture and education have to outside influences such as political ideology. Furthermore, how a participant interprets particular words in the items is of importance to consider. Therefore, we intend to interview a group of PSTs to ascertain their interpretations of particular items. By doing so, we will be able to refine the wording of certain items that were dropped and/or add additional items. In either case, the goal will be to load the items on to the specific factors we have (and not to generate new factors).

The interviews with the PSTs will also help us reframe items 36, 14 and 20 that loaded on multiple factors. In each case, deleting the item from the factor to which it was assigned (based on the higher load) would reduce the alpha. Therefore, a decision was made to retain each item and refine the wording so that it was more likely to load on the assigned factor. Further, by removing item 20 the LM factor would consist of only two items and not meet our criteria for retaining factors. Given the goal of MEELS is to examine PSTs' beliefs about the mathematics education of ELs, having a subscale related to the beliefs about the interconnection of language and mathematics is important.

In the next iteration of MEELS, once the items are refined, we intend to also carry out Confirmatory Factor Analysis (CFA) and test-retest reliability. Confirmatory Factor Analysis (CFA) allows us to test the five factors that we have. The test-retest reliability ensures that the instrument has temporal stability. Specifically we will administer MEELS to a group of PSTs at differing points in the semester (but no longer than a few weeks of each other) to determine if there is a high correlation between their scores at each point. The purpose of administering MEELS at multiple points within a short period of time is to speak to stability in the given constructs.

Finally, our long-term goal is to establish predictive validity of the MEELS. This would require tracking the PSTs into their teaching careers to establish a relationship between the PSTs performance on the MEELS and their teaching of ELs in their mathematics classes. The latter would require the development of observational instruments that could measure the performance of the teacher in implementing best practices for teaching ELs. Currently, we are not aware of such an instrument.

Implications

The Mathematics Education of English Learners Scale (MEELS) is a powerful and necessary instrument for teacher preparation programs. At the end of our validation and reliability process, we will have developed a measure that can be used to ascertain the beliefs of PSTs regarding the mathematics education of ELs on a large scale, an undertaking that we have not come across in our review of the literature. Additionally, among other things, MEELS can be shared with individuals in the field so that the effectiveness of particular interventions with ELs can be measured.

While still in the early stages of reliability and validity testing for our instrument, we have generated some important implications for the field of mathematics education, in general, and teacher education, specifically. Though the content validity of MEELS was determined and only a few (about 4%) participants reported any vague or confusing questions, our initial analysis revealed that not all of the items we developed mapped onto our initial dimensions. In other words, some of the items in MEELS did not describe the beliefs with which the wording appeared. Thus the use of surveys that are not shown to be statistically validated and reliable may be problematic, as they may not always be measuring what they appear to be on the surface.

References

Burns, R. B., & Burns, R. A. (2008). *Business research methods and statistics using SPSS*. Thousand Oaks, CA: Sage.

Byrnes, D. A., & Kiger, G. (1994). Language attitudes of teachers scale (LATS). *Educational and Psychological Measurement, 54*, 227–231. doi:10.1177/0013164494054001029

Civil, M. (2007). Building on community knowledge: An avenue to equity in mathematics education. In N. Nassir & P. Cobb (Eds.), *Improving access to mathematics: Diversity and equity in the classroom* (pp. 105–117). New York: Teachers College Press.

Cooney, T., Shealy, B., & Arvold, B. (1998). Conceptualizing belief structures of preservice secondary mathematics teachers. *Journal for Research in Mathematics Education, 3,* 306–333. doi:10.2307/749792

Costa, J., McPhail, G., Smith, J., & Brisk, M. E. (2005). Faculty first: The challenge of infusing the teacher education curriculum with scholarship on English language learners. *Journal of Teacher Education, 56*, 104–118. doi:10.1177/002248710427119

Davis, L. E., & Turner, J. S. (1993). *An investigation of the cultural sensitivity level of elementary preservice teacher*. Paper presented at the annual meeting of the Mid-South Educational Research Association. (ERIC Document Reproduction Service No. ED 372 0540)

Echevarría, J., Vogt, M., & Short, D. J. (2008). *Making content comprehensible for English learners: The SIOP model* (2nd ed.). Boston: Pearson.

Gandára, P., & Contreras, F. (2009). *The Latino education crisis: The consequences of failed social policies*. Cambridge, MA: Harvard University Press.

Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, *39*, 32–36. doi:10.1007/BF02291575

Larke, P. J. (1990). Cultural Diversity Awareness Survey Inventory: Assessing the sensitivity of preservice teachers. *Action in Teacher Education*, *12*(3), 23–30.

Lucas, T., & Grinberg, J. (2008). Reading to the linguistic reality of mainstream classrooms: Preparing all teachers to teach English language learners. In M. Cochran-Smith, S. Feiman-Nemser, D. J. McIntyre, & K. E. Demers (Eds.), *Handbook of research on teacher education: Enduring questions in changing contexts* (3rd ed.) (pp. 606–636). New York: Routledge.

McLeman, L., & Fernandes, A. (under review). Regarding the mathematics education of English learners: Clustering the conceptions of prospective teachers from urban universities. Submitted to the *Journal of Urban Mathematics Education*.

Moschkovich, J. N. (2010). Language(s) and learning for mathematics: Resources, challenges, and issues for research. In J. N. Moschkovich (Ed.), *Language and mathematics education: Multiple perspectives and directions for research* (pp. 1–28). Charlotte, NC: Information Age.

Pett, M. A., Lackey, N. R., & Sullivan, J. J. (2003). *Making sense of factor analysis: The use of factor analysis for instrument development in health care research*. Thousand Oaks, CA: Sage.

Philipp, R. A. (2007). Mathematics teachers' beliefs and affect. In F. K. Lester, Jr. (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 257–315). Reston, VA: National Council of Teachers of Mathematics.

Pohan, C. A., & Aguilar, T. E. (2001). Measuring educators' beliefs about diversity in personal and professional contexts. *American Educational Research Journal, 38*, 159–182. doi:10.3102/00028312038001159

Reeves, J. (2006). Secondary teacher attitudes toward including English-language learners in mainstream classrooms. *The Journal of Educational Research, 99*, 131–142. doi:10.3200/JOER.99.3.131-143

Stevens, J. (1992). *Applied multivariate statistics for the social sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Tabachnick, B. G., & Fidell, L. S. (1996). *Using multivariate statistics* (3rd ed.). New York: Harper Collins.

Walker, A., Shafer, J., & Iiams, M. (2004). 'Not in my classroom': Teacher attitudes toward English language learners in the mainstream classroom. *NABE Journal of Research and Practice*, *2,* 130–160. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.131.5682&rep=rep1&type=pdf#page=14