

Development and Validation of the Mathematics Teaching Efficacy Scale: Confirmatory Factor Analysis

Ryang, Dohyoung (UNCG)

I. Introduction

Teacher efficacy refers to a teacher's self-perceived beliefs regarding his or her ability to organize and execute courses of actions to successfully accomplish a specific teaching task in a particular context (Tschannen-Moran, Woolfolk Hoy, & Hoy; 1998). Researchers revealed that teaching efficacy can influence a teacher's instructional effectiveness as well as student motivation and achievement (Barr, 2005; Herman, 2000; Mojavezi & Tamiz, 2012; Nelson, 2007). Since teacher efficacy are subject-matter specific (Bandura, 1997; Tschannen-Moran et al., 1998), mathematics teaching efficacy becomes a powerful factors to teachers' future behavioral actions in teaching mathematics and student mathematical outcome.

Efficacy beliefs are claimed as most suitably measured within the context of specific behaviors (Henson, Kogan & Vacha-Haase, 2001; Pajares, 1996). But, the development of a measure should not be so specific that they lose their predictive power and only address very particular skills or context (Tschannen-Moran et al., 1998). Thus, developing an

instrument in an appropriate level measuring a type of efficacy beliefs has grown as one of major issues in the research studies of teacher efficacy (Tschannen-Morann et al., 1998).

Culture influences teacher's beliefs on self-efficacy. Despite the extensive research on teacher efficacy in Western cultures, a limited number of studies have been made in non-Western cultures like Brazil, Israel, Hong-Kong, South Korea, and Sri Lanka (Ares, Gorrell, & Boakari, 1999; Cheung, 2006; Gorrell, Hazareesingh, Carlson, & Stenmalm-Sjoblom, 1993; Gorrell & Hwang, 1995; Lin & Gorrell, 1999; Rich, Lev, & Fischer, 1996). These studies indicated that preservice teachers from different cultures vary in the degree to which they believe they are effective in their teaching. That is, teacher efficacy is strongly influenced by uniquely cultural variables (Lin, Gorrell, & Taylor, 2002). In addition, differences in mathematics teaching efficacy in different cultures were found. Cakiroglu (2008) compared teaching efficacy of preservice teachers in the U. S. and Turkey. He reported that preservice teachers in Turkey tend to have a stronger belief that teaching can influence student learning when compared with preservice teachers in the U.S.; however, a similar difference was not observed for personal mathematics teaching efficacy.

The purpose of this study is to develop a valid instrument measuring mathematics teaching efficacy beliefs of Korean teachers. However, since efficacy is context-specific, the school level (elementary vs. secondary) in which a teacher is teaching mathematics

* Received March 7, 2014; revised May 2, 2014; accepted May 12, 2014

* ZDM Classification: B59

* MSC2000 Classification : 97C70

* Key Words : efficacy, factor analysis, mathematics, preservice teachers

* The paper was presented at the Spring Conference of the Korean Society of Mathematical Association in Hankuk University of Foreign Studies on April 4-5, 2014.

matters on the teacher's efficacy level. Also, secondary mathematics teachers major in mathematics while elementary teachers instruct all subjects and thus they possibly have different levels of efficacy beliefs in different subjects they teach. So, research interest goes first to elementary teachers rather than secondary teachers. Also, since preservice teachers' efficacy level is changeable during the teacher education program, the present study intends to develop a scale for preservice teachers.

With considering the level of specification (mathematics), cultural disparity (Korea), teacher maturation (preservice), and school level (elementary), this study will give an answer to the question: How to develop a scale to measure mathematics teaching efficacy beliefs of Korean preservice elementary teachers?

II. Theoretical Background

The Mathematics Teaching Efficacy Beliefs Instrument (MTEBI) was validated in Bandura's 2-dimensional conceptualization of self-efficacy (Enoch, Smith & Huinker, 2000). The MTEBI consists of 21 items in the two variables: Personal Mathematics Teaching Efficacy (PMTE) and Mathematics Teaching Outcome Expectancy (MTOE). This instrument has contributed in many research studies on mathematics teaching efficacy in the United States or similar cultures. For example, highly efficacious teachers are more effective mathematics teachers than teachers with a lower sense of efficacy (Swars, 2005); teachers with a high self-efficacy are more willing to be creative and to use inquiry-based methods of teaching mathematical concepts (Wilkins, 2008); mathematics teaching efficacy is positively related to methods courses and field experiences (Evans, 2011; Swars, Smith, Smith, & Hart, 2009; Utley, Bryant, & Moseley, 2005); mathematics teaching efficacy has negative

correlations to mathematics anxiety (Gresham, 2008, Swars, Daane, & Giesen, 2006); preservice teachers' ages, lower division mathematics history, and mathematics methods course performance, had a significant relationship with their mathematics teaching efficacy (Brown, 2012).

In the frame of Bandura's efficacy theory, the first teacher efficacy measure was Gibson and Dembo's (1984) Teacher Efficacy Scale (TES). Using this scale, Gorrell et al. (1993) found that U.S. preservice teachers were more positive in their general teaching efficacy than the Swedish and Sri Lankan teachers, but Sri Lankan pre-service teachers revealed higher levels of personal efficacy than U.S. preservice teachers; the findings however may distort some important conceptual differences in teacher efficacy because they did not consider the factor structures of the instrument.

Furthermore, there had been controversy on the factorial validity of the TES. For example, For example, the 2-factor structure of TES was observed in Rich et al.'s (1996) study in an Israel sample. In contrast, Kushner (1993) argued that the 2-factor model of the TES did not fit the data very well. Other researchers tested various factor models on the TES such as another 2-factor model (Guskey & Passaro, 1994), 3-factor models (Emmer & Hickerman, 1991; Soodak & Podell, 1996; Woolfolk & Hoy, 1990), and a 4-factor model (Lin & Gorrell, 1999). Brouwers and Tomic (2003) tested the aforementioned 2, 3, and 4-factor models on 540 Dutch teachers, and found that the 4-factor model fitted the data better than the others but its fit indices did not reach the recommended level. Finally, the TES is not suitable for a research study anymore.

These controversies among researchers about the TES led researchers to explore new teacher efficacy scales in different perspectives. Tschannen-Moran and Hoy (2001) developed a 24-item instrument, the

Teacher Self-Efficacy Scale, with the three subscales: efficacy for instructional strategies, efficacy for classroom management, and efficacy for student engagement. The alpha reliability of the three subscales and the whole instrument were reported by .87, .91, .90, and .94, respectively. Bandura (2006) also constructed a 30-item instrument with the seven variables: efficacy to influence decision making, efficacy to influence school resources, instructional efficacy, disciplinary efficacy, efficacy to enlist parental involvement, efficacy to enlist community involvement, and efficacy to create a positive school climate. Unfortunately, Bandura's instrument was not experimentally tested.

The MTEBI was developed in a U. S. sample with Bandura's theory, as the TES was. The MTEBI can be regarded as a succession of the TES in the context of mathematics teaching. In fact, many items in the MTEBI were similar in the TES. Despite the usefulness of the MTEBI in recent research studies, the issue of factorial validity of the MTEBI still remains in that the TES is not useful for a research study. So, it is hypothesized that the MTEBI has potential weakness in factorial validity. Though the MTEBI showed an acceptable level of validity, societal shifts for the past 14 years may influence on change in teachers' efficacy beliefs in mathematics teaching.

In addition, the MTEBI is not guaranteed for working in other cultures. There were a few research studies testing the validity the MTEBI in other cultures. Alkhateeb (2004) translated the MTEBI into Arabic and tested on 144 Arabic speaking Jordan undergraduate education majors. The factorial validity was explored by a principal component analysis with varimax rotation. The alpha coefficients were .84 for the PMTE and .75 for the MTOE. In different vein, Chang (2003) translated the MTEBI into traditional Chinese and pilot tested on the sample of 30 Taiwanese senior students. The Chinese version

MTEBI had the coefficient alpha .87 for the PMTE and .67 for the MTOE. Cakiroglu (2008) pilot tested a Turkish MTEBI and reported alpha coefficient .77 for the PMTE and .65 for the MTOE. Because of the lower alpha values of MTOE subscales in Chinese and Turkish studies, the results should be minimally accepted.

Recently, Ryang (2013a, 2013b) modified the MTEBI for Korean elementary and secondary preservice teachers so factorial validity was established on each version. However, the items of these revised versions were restricted within the American MTEBI so they do not fully reflect original values and concepts shared among teachers, teacher educators, researchers, and educational policy makers of South Korea. In order to expand discussion on mathematics teaching efficacy to Korean preservice teachers, a new instrument needs to be developed.

III. Methods

This study develops a mathematics teaching efficacy measure within the frame of Bandura's self-efficacy theory. Initially, the fifty eight items were developed from the literature review and Korean mathematics teacher educators' suggestion. The study focuses on establishing the factorial validity of the instrument. Inappropriate items are deleted from the first form of 58 items.

1. Construct and Variables

According to Bandura (1997), self-efficacy is a cue from social behaviors, personal cognitive interpretations, and environmental influences that intertwine interactively. Thus, perceptions determine resultant action consequences (triadic reciprocal determinism). And, self-efficacy is better conceptualized as one of the two dimensions, the personal efficacy (PE) and the outcome expectancy

(OE) in a belief system, which intervening stimuli and responses (situational-interaction).

The construct to be measured in the present study is preservice teachers' mathematics teaching efficacy. In Bandura's perspective, the variables explaining this construct are the PE and the OE in mathematics teaching. Thus, the instrument should have the two subscales dealing with these two variables, named the Mathematics Teaching Personal Efficacy (MTPE) and the Mathematics Teaching Outcome Expectancy (MTOE). The MTPE was formerly named Personal Mathematics Teaching Efficacy (PMTE) in Enoch et al.'s (2000) study.

The MTPE items deal with personal effectiveness in mathematics teaching; these are stated in the first person with the future tense since they *will* teach mathematics in the future. The MTOE items describe general beliefs about the effect of the social and cultural sources on student outcome for mathematics teaching. These are stated in the third person with the present tense. For illustrative purpose, an item in each variable is shown below.

| <i>Variable</i> | <i>Item</i> |
|-----------------|---|
| MTPE | I will not be able to teach mathematics effectively. |
| MTOE | Teachers are the most powerful factor to student mathematics achievement. |

2. Participants

The total participants were 1015 Korean elementary preservice teachers enrolled in seven out of 12 national universities of education in South Korea. Among them, 688 (68.5%) were female and 317 (31.5%) male; 147 (14.6%) were freshmen, 403 (40.1%) sophomores, 291 (29.0%) juniors, and 164 (16.3%) seniors. Average age was 22.59 years old with standard deviation 3.311.

3. The First Form

Among the initial form of 58 items, the first 21 items were a modified copy of the MTEBI. And, some MTEBI items were revised into items in the first form. For example, I18 in the MTEBI is that: *Given a choice, I will not invite the principal to evaluate my mathematics teaching.* However, this item strayed away from Korean school culture in that relationship between a teacher and the school principal is much more hierarchical than in the U. S. schools. Also, Korean schools implement 'lesson study' so most teachers have no hesitate to open their teaching to others, which may be an opposite stance to American teachers' independence in their classroom, so called, 'teacher isolation.'

Korean experts in mathematics teacher education recommended including some items in the first form. Those items were similar to the MTEBI items but simpler and smoothly worded. For example, see Item 26 (*I certainly will teach mathematics well in an open class*) and Item 29 (*I will teach mathematics in such a way that the students easily understand the concept*). Also, the first form included items from literature such as other efficacy instrument. For example, Item 33 (*A teacher's effectiveness in mathematics teaching has little influence on the mathematics achievement of students with low motivation*) is a modification of an item in Gibson and Dembo's scale. Item 43 (*I will be able to implement an innovative mathematics teaching strategies*) is modified from an item in Tschannen-Moran and Hoy's scale (2001).

In addition, some items were reflecting current trends in mathematics education of Korea like student-centered activities, differentiated or individualized education, and subject-integrated education. For instance, see Item 50 (*I will have a difficulty in adjusting mathematics lessons to the proper level for individual students*), Item 52 (*If a teacher gets students to work mathematical tasks*

together, then their mathematical achievement increases), and Item 54 (*A teachers' use of non-mathematical knowledge in mathematics teaching helps students understand the mathematical concepts*).

4. Data Collection and Analysis

Participants were asked to rate their feelings in the survey items. The survey packet was distributed in a regular class, and administered for 30 minutes. An informed consent form was provided to the program coordinators or the department chairs before distributing the survey. The consent information for the survey participants were provided within the survey packet. The preservice teachers agreed participating in the study and then responded to the survey questions. Ninety six participants did not respond to at least one item; these cases were list-wise deleted so the valid data set was constituted from 919 participants.

The instrument uses a 5-point rating scale: Strongly Disagree, Disagree, Uncertain, Agree, and Strongly Agree. The items were coded by the letter P for the MTPE items or the letter O for the MTOE items with the item numbers (e.g., O1, P2), in order to conveniently see the factor structure. Then, the responses were coded by one of five points, from 1 for Strongly Disagree to 5 for Strongly Agree. Negatively worded items were reversely coded (1 = 5, 2 = 4, 4 = 2, 5 = 1).

Appropriate statistical approaches are conducted on the data set. First, the normality of an item variable will be tested. Then, the whole data set ($N = 919$) is divided by the two subsets for different methods of factor analysis. The set was arbitrarily divided considering the gender ratio and the class level ratio. The factor structure is *explored* in the first data set ($N = 419$), and the structure is then *confirmed* in the second data set ($N = 500$). Testing the factor structure in two different sets will increase

cross-validity of the instrument. This study used the IBM SPSS 21 program for the normality test and the exploratory factor analysis and the LISREL 8.80 program for the confirmatory factor analysis.

IV. Results and Discussion

1. Normality Test

Normality is violated by skewness and kurtosis. It is doable using the combined scores of the skewness and kurtosis rather than using the scores separately. The LISREL reported the Z -scores for the skewness and kurtosis combination and their p -values for all 58 items. In the .05 significance level, the 25 MTPE items and the 10 MTOE, together 35 items, items passed the normality test; that is, they have p -values greater than .05. [Table 1] shows these 35 items. Note that four O-items (O7, O35, O36, O45) have means less than 3.

2. Exploratory Factor Analysis

Using the principal component analysis (PCA), the 2-factor structure on the 35 items was explored in the first data set ($N = 419$). The MTPE and the MTOE is hypothesized to have a 1-factor structure, and thus the combined scale of the two subscales will have a 2-factor structure.

1) MTPE scale

The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy index for the 25 MTPE items was .912. Bartlett's test of sphericity was significant ($\chi^2 = 3631.423$, $DF = 300$; $p < 0.001$). The PCA with varimax rotation on the 25 MTPE items extracted 5 components with eigenvalues greater than 1. But, the first component had a distinctively high eigenvalue and others made a smooth decrease of eigenvalues. A 1-factor structure was suggested on the 25 P-items.

[Table 1] Mean, standard deviation, and *Z*-scores of the skewness-kurtosis combination and their *p*-value for the 35 items

| Item | Mean | S. D. | <i>Z</i> -score | <i>p</i> -value |
|------|-------|-------|-----------------|-----------------|
| P6 | 3.721 | 0.798 | 2.325 | 0.313 |
| P8 | 3.976 | 0.782 | 5.355 | 0.069 |
| P11 | 3.399 | 0.856 | 0.049 | 0.976 |
| P15 | 3.621 | 0.870 | 3.554 | 0.169 |
| P16 | 3.623 | 0.853 | 1.117 | 0.572 |
| P18 | 3.375 | 0.928 | 2.278 | 0.32 |
| P20 | 3.723 | 0.841 | 4.166 | 0.125 |
| P21 | 3.585 | 0.861 | 2.608 | 0.271 |
| P22 | 3.329 | 0.813 | 1.006 | 0.605 |
| P23 | 3.432 | 0.887 | 0.964 | 0.618 |
| P24 | 3.430 | 0.982 | 3.268 | 0.195 |
| P26 | 3.224 | 0.873 | 1.788 | 0.409 |
| P29 | 3.714 | 0.712 | 3.245 | 0.197 |
| P31 | 3.601 | 0.752 | 1.341 | 0.511 |
| P32 | 3.403 | 0.865 | 0.271 | 0.873 |
| P38 | 3.754 | 0.760 | 4.229 | 0.121 |
| P40 | 3.449 | 0.930 | 2.977 | 0.226 |
| P43 | 3.432 | 0.730 | 1.246 | 0.536 |
| P47 | 3.778 | 0.709 | 5.163 | 0.076 |
| P48 | 3.363 | 0.746 | 0.243 | 0.886 |
| P50 | 3.148 | 0.947 | 1.976 | 0.372 |
| P53 | 3.141 | 0.854 | 0.161 | 0.923 |
| P55 | 3.494 | 0.796 | 0.689 | 0.709 |
| P56 | 3.332 | 0.967 | 2.374 | 0.305 |
| P58 | 3.606 | 0.966 | 3.08 | 0.214 |
| O1 | 3.454 | 0.853 | 4.599 | 0.100 |
| O7 | 2.878 | 0.898 | 0.751 | 0.687 |
| O25 | 3.656 | 0.780 | 5.161 | 0.076 |
| O35 | 2.902 | 0.996 | 3.012 | 0.222 |
| O36 | 2.320 | 0.930 | 5.24 | 0.073 |
| O42 | 3.294 | 0.950 | 1.751 | 0.417 |
| O44 | 3.582 | 0.764 | 3.012 | 0.222 |
| O45 | 2.294 | 0.873 | 5.302 | 0.071 |
| O52 | 3.406 | 0.840 | 2.307 | 0.316 |
| O54 | 3.709 | 0.839 | 1.944 | 0.378 |

2) MTOE scale

The KMO measure of sampling adequacy index was .758. Bartlett's test of sphericity was significant ($\chi^2 = 635.730$, $DF = 45$, $p < 0.001$). The PCA with varimax rotation on the 10 MTOE items initially extracted 2 components with eigenvalues greater than 1. The scree plot suggested that the scale have a 2-factor structure. A 2-factor solution on the 10 MTOE items indicated that the three items (O35, O36, O45) constituted a minor factor. After deleting these items, the remaining seven O-items were suggested to have a 1-factor structure.

3) Combined scale

The 25 P-items and the remaining seven O-items were combined to make a scale of the 32 items. The KMO index was .909; Bartlett's sphericity test was significant ($\chi^2 = 4440.460$, $DF = 496$, $p < 0.001$). The PCA with varimax rotation on this model extracted 7 components with eigenvalues greater than 1. But, first two components had distinctively higher eigenvalues than the others. A 2-factor structure was suggested. Consecutively, the PCA with varimax rotation extracted 2-factor solution on the 32 items. While all O-items were loaded to the component 2, two items (P40, P53) were loaded to the component 2 and seven items (P29, P31, P32, P38, P43, P47, P55) loaded to both components. After deleting these nine P-items, the 23 items (16 P-items and 7 O-items) tremendously increase factorial validity. [Table 2] showed the 2-factor solution PCA on the 23-item model. Factor loading less than .32 (Osborne & Costello, 2005) were erased in the table. The component 1 explained 23.23% and the component 2 explained 11.77% of the total variance; with together 35% of the total variance. By reading the pattern of the factor loadings in [Table 2], the component 1 and the component 2 must be the MTPE and the MTOE, respectively.

[Table 2] PCA with Varimax Rotation on the 23 Items

| Item | Component | |
|------|-----------|------|
| | 1 | 2 |
| P23 | .697 | |
| P16 | .690 | |
| P22 | .679 | |
| P24 | .653 | |
| P21 | .639 | |
| P11 | .629 | |
| P26 | .607 | |
| P58 | .599 | |
| P8 | .596 | |
| P56 | .551 | |
| P20 | .517 | |
| P15 | .490 | |
| P18 | .480 | |
| P6 | .417 | |
| P50 | .354 | |
| P48 | .330 | |
| O42 | | .730 |
| O25 | | .630 |
| O44 | | .619 |
| O52 | | .553 |
| O1 | | .538 |
| O7 | | .526 |
| O54 | | .411 |

3. Reliability Analysis

In the 23-item model, the 16 P-items had reliability, Cronbach's internal consistency $\alpha = .865$, the seven O-items reliability $\alpha = .686$, and the global scale reliability $\alpha = .860$. There perhaps exists an item that reduces the reliability of the scale which the item belongs to. To find a weak item, alpha reliability after deleting an item was compared with that before deleting the item. Deleting an item, except the two items (P50, O7), did not exceed the scale reliabilities

before deleting the item. Deleting P50 did not increase the MTPE scale reliability; deleting O7 does not increase the global reliability. These two items were not immediately deleted but flagged for further investigation.

4. Confirmatory Factor Analysis

Structural Equation Modeling (SEM) is a useful way to confirm the factorial structure on the Likert-type responded data set. The 2-factor model of the 23-item scale was *confirmed* using the SEM. The LISREL reported various fit indices indicating how the conceptual 2-factor model is reasonably fitting the empirical data.

The fit indices used in this study are the ratio of χ^2 to the degree of freedom (DF), root mean square error of approximation (RMSEA), standardized root mean square residual (SRMR), comparative fit index (CFI), and goodness-of-fit index (GFI). When deleting an item one at a time, these indices would be improved. The target values of these indices are $\chi^2/DF = 3$, (Abell et al., 2009), RMSEA = .05, SRMR = .05, CFI = .90, and GFI = .90 (Bae, 2006). The calculated fit indices for the 23-item model were calculated on the second data set ($N = 500$); see [Table 3]. All of fit indices did not reach the target value stated above, indicating that the 23-item model should be modified by deleting weak items.

[Table 3] Some Models and Their Fit Indices

| Fit Index | 23-Item Model | 18-Item Model A | 18-Item Model B |
|-------------|---------------|-----------------|-----------------|
| χ^2 | 822.08 | 343.52 | 268.44 |
| DF | 229 | 134 | 131 |
| χ^2/DF | 3.59 | 2.566 | 2.05 |
| RMSEA | .072 | .056 | .046 |
| SRMR | .061 | .049 | .044 |
| CFI | .78 | .85 | .90 |
| GFI | .87 | .93 | .95 |

High covariance errors existed between an item and the opponent variable or between items in the opponent variables (e.g., P6 → MTOE, O4 → MTPE, P15 ↔ O7) that should not be ignored. Those items were deleted one at a time in the order of the most χ^2 -decrease. After deleting such five items (O7, P48, P56, P22, P23), the 18 items (Model A) had the fit indices: $\chi^2/DF = 2.564$, RMSEA = .056, SRMR = .049, CFI = .85, GFI = .93. RMSEA and CFI did not reach the target values yet. Though no more covariance errors existed cross the variables, tiny covariance error still exists within some P-items (P18 ↔ P24; P20 ↔ P16 ↔ P11). Existence of small covariance error in the same variable is not unusual (Enoch et al., 2000). By allowing those errors increases the model fit indices. The 18-item model with allowing those three covariance errors (Model B) had fit-indices; $\chi^2/DF = 2.049$, RMSEA = .046, SRMR = .044, CFI = .90, GFI = .95, which reached the target values and even showed an excellent level of the model fit. Note that Items O7 and P50 were flagged in reliability analysis. Here, P50 survived while O7 was deleted. Thus, the items of the mean less than 3 were all removed from the instrument.

The SEM also examines the linear regression model of each dependent variable (item) to a latent variable (MTPE or MTOE). The regression coefficients serve as *structural* factor loadings which provide more rigorous values than (*exploratory*) factor loadings in a PCA solution. The LISREL provides the structural coefficient estimates, λ (factor loading) and ϕ (correlation between the latent variables) with their standard errors. Rather than λ itself, the t -value was used to determine the item's significance. In the significance level of .05, an item of which $|t| < 1.96$ should be deleted. The LISREL reported λ and t for every 18 item; see [Table 4]. All items passed this criteria.

[Table 4] LISREL Estimates (Maximum Likelihood) of the 18-Item Model

| Item | MTPE | | | MTOE | | |
|------|-----------|-------|-------|-----------|-------|------|
| | λ | Error | t | λ | Error | t |
| P6 | .33 | .04 | .70 | | | |
| P8 | .40 | .03 | 12.69 | | | |
| P11 | .40 | .03 | 11.66 | | | |
| P15 | .28 | .04 | 7.19 | | | |
| P16 | .39 | .03 | 11.55 | | | |
| P18 | .40 | .04 | 9.21 | | | |
| P20 | .31 | .04 | 7.57 | | | |
| P21 | .41 | .04 | 11.22 | | | |
| P24 | .58 | .04 | 14.98 | | | |
| P26 | .46 | .03 | 15.08 | | | |
| P50 | .36 | .04 | 9.10 | | | |
| P58 | .41 | .04 | 11.61 | | | |
| O1 | | | | .30 | .05 | 6.47 |
| O25 | | | | .32 | .04 | 7.45 |
| O42 | | | | .37 | .05 | 7.46 |
| O44 | | | | .36 | .04 | 8.17 |
| O52 | | | | .25 | .05 | 5.01 |
| O54 | | | | .10 | .05 | 2.21 |

The factorial construct validity was well established in the 18-item model in two ways: convergence within a group of items and discrimination between the groups of items. The convergence validity was explored by the PCA on the first data set ($N = 419$) where all P-items are loaded to the MTPE and all O-items to the MTOE; see [Table 2]. Further, the convergence was confirmed in the second data set ($N = 500$) where all structural coefficients λ were significant; see [Table 4]. To see discrimination between the MTPE and the MTOE, test the null hypothesis that the two variables covariate completely (Bae, 2006). In the 18-item model, the ϕ -coefficient between the MTPE and the MTOE was .31 with

standard error 0.06. The 99% confidence interval was $0.31 \pm 2.58*(0.06) = (0.155, 0.465)$, which did not include 1. So, the null hypothesis was rejected and it was concluded that the 18-item model has the discriminant construct validity. This model was accepted as a precise instrument and named the *Mathematics Teaching Efficacy Scale* (MTES). The 18 items after renumbering were shown in Appendix.

On the other hand, reliability is another factor that should be considered in developing a research instrument. Previous studies reported the reliabilities of the MTPE and the MTOE; see [Table 5]. The MTPE reliabilities are higher than the MTOE reliabilities in the studies. The MTOE reliabilities in Chang (2003), Cakiroglu (2008), and the present studies were less than .7. These results imply that the MTOE scale is not uniquely fixed cross cultures. In addition, the complexity of current Korean society may little make the influence of the outcome expectancy to teacher efficacy in mathematics teaching. It is here recommended that the MTOE subscale must be minimally used in a research study.

[Table 5] Alpha Reliabilities in Different Cultures

| Researcher (year) | Country | N | MTPE | | MTOE | |
|------------------------|---------|-----|------|-----|------|-----|
| | | | n | a | n | a |
| Enochs et al (2000) | U.S.A. | 324 | 13 | .88 | 8 | .77 |
| Chang (2003) | Taiwan | 30 | 13 | .87 | 8 | .67 |
| Alkhateeb (2004) | Jordan | 144 | 13 | .84 | 8 | .75 |
| Cakiroglu (2008) | Turkey | NR | 13 | .77 | 8 | .65 |
| Ryang (2014)* | Korea | 419 | 12 | .83 | 6 | .68 |

Note. N = Sample size; n = Number of items; NR = Not Reported

*The present study

5. Discussion on Items

The MTES was developed from the item pool including the MTEBI and revised items from the MTEBI. The MTES items were discussed here in comparison to the MTEBI items.

1) MTPE Items

The MTES includes 12 P-items among which eight items were transmitted from the MTEBI. That is, five P-items (I2, I3, I5, I17, I19) in the MTEBI were removed in the new MTES. Those items actually had problems of multiple meaning, awkwardness, tense disagreement, and vagueness. See Ryang, Thompson and Shwery's (2011) work for a detailed account. Instead of those problematic items, newly developed items were believed as having better sense to Korean preservice teachers. So, the new MTES are expected to be more appropriate to the Koreans than the MTEBI.

While the MTEBI has only one item related to opening a classroom, the MTES has couple of items dealing with the matter of opening the classroom; Item 6 (*agree to open class*), Item 9 (*fear to open class*), Item 10 (*class to the public*). See the Appendix for the full statements of the items. The related MTEBI item was that: *Given a choice, I will not invite the principal to evaluate my mathematics teaching*. As mentioned in Section III. 3, this item was regarded as one of the most problematic items to Korean mathematics teacher educator. The item seems to weigh on a teacher's *choice* rather than asking *feeling* of ability about mathematics teaching (Ryang, Thompson, & Shwery, 2011). Thus, instead of the original item (*given a choice*), an alternative (Item 6) was used in this study. Korean preservice teachers in most cases do 'lesson study' in a method course and/or they have experience of participating in a lesson when they were school students. Therefore, the Korean preservice teachers well understand the

openness of a classroom to others as a basic component in the school-based lesson study. An item about opening the class is well-set in preservice teachers' and teacher educators' minds.

Item 11 in the MTES mentions individualized or differentiated teaching in a mathematics classroom. The current trend of mathematics education in Korea emphasizes differentiated education. The trend is well introduced in the teacher education program so Korean preservice teachers apparently understand the importance of this theme. Therefore, the participants of this study consistently responded to the survey so the item survives in the MTES as a statistically solid item.

In the process of factor analysis, 13 P-items were removed from the item pool. Among them, many items were not appropriately stated. Though such an item were written in the format of P-items (first person, future tense), the item is interpretable as describing outcome expectancy. In the first form, P47 (*I will get students to believe they can do well in mathematics*) and P55 (*I will succeed to motivate students low-achieving in mathematics*) are of this case. It is in question whether the ability to get students to believe themselves doing well and/or to motivate students is at source from personal mathematics teaching or at other sources of preservice teachers. These two items may more appeal with the O-item form (If-then, present). For instance: *If I teach mathematics effectively, then students show more interest in doing mathematics.*

Similarly, P40 (*When I really try hard, I can make most unmotivated students improve in mathematics assessment*) are not written appropriately. Regardless of appropriateness of its content, the item is written in the O-item format. Rather, this item is suggested to be stated as a P-item by: *I will be able to make most unmotivated students understand mathematics concept*, or, as an O-item, *If I really try hard, even*

most unmotivated students improve in mathematics assessment. In sum, the vagueness in the context of those three items resulted in converging to the incorrect factor.

2) MTOE Items

When exploring the 1-factor structure of the MTOE, Items O35, O36, and O45 were deleted since they constituted a minor factor so they hurt the 1-factor structure. Those items differs from the other O-items in the MTES. First, these are all negatively worded (*cannot be; cannot help; may not reach*) while the survival O-items are all positively worded (see Appendix). Secondly, these items ask a teacher's responsibility to *all* students' success in mathematics while the other O-items state a teacher's performance and characteristics in a neutral tone without mentioning *some* or *all*. Lastly, these three items are all positively skewed. Since the participants were preservice teachers and want to become an actual classroom teacher, they are assumed to respond to each item with a positive sense. In fact, most items were little bit negatively skewed. The other positively skewed item O7 was deleted when modifying the model by the SEM. So, O-items with negative wording were removed from the instrument. It is noteworthy that OE items are conceptualized as a form: If effective teaching, then positive change in student outcome. Korean preservice teacher were responded as conceptually expected.

Despite the well establishing the factorial validity on the MTES, a further study will test O54: *A teacher's use of non-mathematical knowledge in mathematics teaching helps students understand the mathematical concept.* This item had a low factor loading ($\lambda = .1$, $t = 2.21$) on the second data set; see [Table 4]. The t -value is meaningful with the level of significance .05 but it is not meaningful with significance level .02. However, this item deals with

use of other subject materials in teaching mathematics which reflects an educational theme currently emphasized in mathematics education of South Korea. In current status, Item O54 remains in the MTES as Item 18.

Item 17 in the MTES reflects constructivist conception on teaching and learning. In the item, a teacher is a helper or facilitator for students to make a group and work together. Modern education differs from traditional education in that a student is person to construct knowledge. Preservice teachers learn and establish constructivist conceptions on teaching and learning during the teacher education program. This item sets value in the MTES applying to a current conceptual venue of teacher education in Korea.

Other O-items in the MTES are similarly stated to an item in the MTEBI. But, an O-item in the MTES is clearly stated while the corresponding MTEBI item is not. For example, Item 14 (*A student's lack of mathematical knowledge and attitudes can be overcome by good teaching*) in the MTES corresponds to I9 (*The inadequacy of a student's mathematics background can be overcome by good teaching*) in the MTEBI. While 'inadequacy of mathematics background' in the MTEBI is vague in its meaning to Korean preservice teachers, 'lack of mathematics knowledge and attitudes' is much clearer so the item obtain converging responses from Korean preservice teachers.

Item 13 (*When a student does better than usual in mathematics, it is because the teacher exerted extra effort*), the first O-item in the MTES, is the only survival from the MTEBI. That is, one out of eight O-items in the MTEBI remained in the MTES; survival rate was $1/9 = .11$. This item is most commonly accepted to mathematics teachers from different cultures. On the other hand, teachers' expectancy to student mathematical outcome may be largely influenced by the culture they reside in.

V. Conclusions and Implication

Mathematics teaching efficacy is an important construct to predict a teacher's future behavioral actions in mathematics teaching. It is central to develop an appropriate scale measuring the level of mathematics teaching efficacy beliefs. An existing scale MTEBI is considered as appropriate for U. S. preservice teachers, but it may not be appropriate for Korean preservice teachers due to the difference in cultural context between the two countries. This study focused on establishing the factorial validity of the MTES. In brief, the PCA explored the 2-factor structure of the MTES, and then the SEM confirmed the structure. Thus, the MTES has the two subscales, the 12-item MTPE and the 6-item MTOE. The MTES is more precise than the MTEBI for Korean elementary preservice teachers. It will contribute more to improve the teacher education program in Korea.

The article discussed the items in the MTES especially in comparison to the corresponding items in the MTEBI and other items. In sum, eight items of the 12 P-items in the MTES came from the MTEBI P-items while only one MTEBI O-item survives in the MTES. With noting that personal teaching efficacy is the sense of an individual teacher's own ability while outcome expectancy is the feeling on students' achievement in learning, the result implies that Korean preservice teachers' personal efficacy in mathematics teaching is more relying on a person's inner attributions to mathematics than outcome expectancy. In contrast, outcome expectancy relies more on general conceptions and beliefs surrounding educational settings in South Korea. The MTPE can become a unified construct for all with no regard of teachers' ethnical and cultural background while the MTOE varies culture to culture.

The reliabilities in [Table 5] are also support the same trait of the mathematics teaching efficacy beliefs.

The MTOE reliabilities are lower than the MTPE reliabilities in various cultures. And, most cases, the MTOE reliabilities are lower than .7, implying that the MTOE is not consistent cross cultures. Also, this interpretation coincides with Cakiroglu's (2008) finding that Turkish preservice teachers tend to have a stronger belief that teaching can influence student learning than American preservice teachers; however, a similar difference was not observed for personal mathematics teaching efficacy.

In addition, the MTES reflects recent educational trends in mathematics curriculum of South Korea such as individualized and/or differentiated teaching, student-centered teaching, and subject-integrated teaching. Also, the items in the MTES are more clearly stated than the items in the MTEBI. The robust stance of the MTES will make internal as well as, with hope, international research on mathematics teaching efficacy. Thus, the results of research studies on mathematics teaching efficacy will give many hints and information to mathematics teacher educators in reforming the teacher preparation program.

Finally, teachers' self-perceived efficacy would be changed as the society shifts over a time period. And teacher efficacy also varies spatially culture to culture. In this sense, the validity of an instrument measuring a type of self-efficacy of teachers is not invariant but a property of ongoing process. Researchers ought to alert this matter continually in the future.

References

- Abell, N., Springer, D. W., & Kamata, A. (2009). *Developing and validating rapid assessment instruments*. Oxford University Press.
- Alkhateeb, H. M. (2004). Internal consistency reliability and validity of the Arabic translation of the mathematics teaching efficacy beliefs instrument. *Psychological Reports, 94*, 833-838.
- Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin, 103*, 411-423.
- Ares, N., Gorrell, J., & Baokari, F. (1999). Expanding notions of teacher efficacy: A study of preservice education students in Brazil. *Journal of Interdisciplinary Education, 3*, 1-28.
- Bae, B. (2006). *LISREL structural equation model: Understanding, practice, and programming*. Seoul: Chongram.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: W. H. Freeman.
- Bandura, A. (2006). Guide for constructing self-efficacy scales. In F. Pajares & T. Urdan (Eds.), *Self-efficacy beliefs of adolescents* (pp. 307-337). Charlotte, NC: Information Age Publishing.
- Barr, L. C. (2005). *The effects of parent and teacher self-efficacy on the academic achievement motivation of preschool children*. Unpublished master's thesis, Wake Forest University.
- Brouwers, A., & Tomic, W. (2003). A test of the factorial validity of the Teacher Efficacy Scale. *Research in Education, 69*, 67-79.
- Brown, A. B. (2012). Non-Traditional Preservice Teachers and Their Mathematics Efficacy Beliefs. *School Science and Mathematics, 1123*, 191-198.
- Cakiroglu, E. (2008). The teaching efficacy beliefs of pre-service teachers in the USA and Turkey. *Journal of Education for Teaching, 34*(1), 33-44.
- Chang, Y. L. (2003). *An examination of knowledge assessment and self-efficacy ratings in teacher preparation programs in Taiwan and the United States*. Unpublished Doctoral Dissertation, University of Idaho.
- Cheung, H. Y. (2006). The measurement of teacher efficacy: Hong-Kong primary in-service teachers. *Journal of Education for Teaching, 32*, 435-451.

- Emmer, E. T., & Hickman, J. (1991). Teacher efficacy in classroom management and discipline. *Educational and Psychological Measurement, 51*, 755-65.
- Enochs, Smith, & Huinker (2000). Establishing factorial validity of the mathematics teaching efficacy beliefs instrument. *School Science and Mathematics, 100*(4), 194-202.
- Evans, B. R. (2011). Content knowledge, attitudes, and self-efficacy in the mathematics New York City Teaching Fellows (NYCTF) program. *School Science and Mathematics, 111*(5), 225-235.
- Gibson, S., & Dembo, M. H. (1984). Teacher efficacy: A construct validation. *Journal of Educational Psychology, 76*, 569-582.
- Gorrell, J., Hazareesingh, N. A., Carlson, H. L., & Stenmalm-Sjoblom, L. (1993, August). A comparison of efficacy beliefs among preservice teachers in the United States, Sweden and Sri Lanka. Paper presented at the meeting of the American Psychological Association, in Toronto, Canada, August 20-24.
- Gorrell, J., & Hwang, Y. S. (1995). A study of self-efficacy beliefs among pre-service teachers in Korea. *Journal of Research and Development in Education, 28*, 101-105.
- Gresham, G. (2008). Mathematics anxiety and mathematics teacher efficacy in elementary preservice teachers. *Teaching Education, 19*(3), 171-184.
- Guskey, T. R. & Passaro, P. D. (1994). Teacher efficacy: a study of construct dimensions. *American Educational Research Journal 31*, 627-643.
- Henson, R. K., Kogan, L. R., Vacha-Hasse, T. (2001). A Reliability generalization study of the teacher efficacy scale and related instruments. *Educational and Psychological Measurement, 61*(3), 404-420.
- Herman, P. (2000). *Teacher experience and teacher efficacy: Relations to student motivation and achievement*. Unpublished doctoral dissertation. University of North Carolina at Chapel Hill.
- Kushner, S. N. (1993, February). Teacher efficacy and preservice teachers: A construct validation. Paper presented at the sixteenth annual meeting of the Eastern Educational Research Association, Clearwater Beach FL (ERIC Document Reproduction Service No. 356265)
- Lin, H., & Gorrell, J. (1999). Exploratory analysis of preservice teacher efficacy in Taiwan. *Teaching and Teacher Education: An International Journal of Research and Studies, 17*(5), 623-635.
- Lin, H., & Gorrell, J. (2002). Influence of culture and education on U.S. and Taiwan preservice teachers' efficacy beliefs. *Journal of Educational Research, 96*(1), 37-46.
- Mojavezi, A. & Tamiz, M. P. (2012). The Impact of Teacher Self-efficacy on the Students' Motivation and Achievement. *Theory and Practice in Language Studies, 2*(3), 483-491.
- Nelson, S. L. (2007). *Teacher efficacy and student motivation: A link to achievement in elementary mathematics*. Unpublished doctoral dissertation, University of South Dakota.
- Osborne, J. W., & Costello, A. B. (2005). Best practices in exploratory factor analysis: Four recommendations forgetting the most from your analysis. (7),1-9.
- Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Review of Educational Research, 66*, 543-578.
- Rich, Y., Lev, S., & Fischer, S. (1996). Extending the concept and assessment of teacher efficacy. *Educational and Psychological Measurement 56*, 1015-1025.
- Ryang, D. (2013a). Developing the Mathematics Teaching Efficacy Beliefs Instrument Korean version for secondary prospective mathematics teachers. *Journal of the Korean Society of*

- Mathematical Education Series A: The Mathematical Education, 52*(2), 231-245.
- Ryang, D. (2013b). Development of the Mathematics Teaching Efficacy Beliefs Instrument Korean version for elementary preservice teachers. *Journal of the Korean Society of Mathematical Education Series A: The Mathematical Education, 52*(3), 363-377.
- Ryang, D., Thompson, T., & Shwery, C. (2011). Analysis of Korean mathematics teacher educators' response to the Mathematics Teaching Efficacy Beliefs Instrument. *Journal of the Korean Society of Mathematical Education Series D: Research in Mathematical Education, 15*(3), 229-250.
- Soodak, L., & Podell, D. M. (1996). Teacher efficacy: Toward understanding of a multifaceted construct. *Teaching and Teacher Education, 12* (4), 401-411.
- Swars, S. L. (2005). Examining perceptions of mathematics teaching effectiveness among elementary preservice teachers with differing levels of mathematics teacher efficacy. *Journal of Instructional Psychology, 32*(2), 139-147.
- Swars, S. L., Daane, C. J., & Giesen, J. (2006). Mathematics Anxiety and Mathematics Teacher Efficacy: What is the Relationship in Elementary Preservice Teachers? *School Science and Mathematics, 106*(7), 306-315.
- Swars, S. L., Smith, S. Z., Smith, M. E., & Hart, L. C. (2009). A longitudinal study of effects of a developmental teacher preparation program on elementary prospective teachers' mathematics beliefs. *Journal of Mathematics Teacher Education, 12*, 47-66.
- Tschannen-Moran, M., & Hoy, A. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and Teacher Education, 17*, 783-805.
- Tschannen-Moran, M., Woolfolk Hoy, A., & Hoy, W. K. (1998). Teacher efficacy: Its meaning and measure. *Review of Educational Research, 68*, 202-248.
- Utley, J., Bryant, R., & Moseley, C. (2005). Relationship between science and mathematics teaching efficacy of preservice elementary teachers. *School Science and Mathematics, 105*(2), 82-87.
- Wilkins, J. L. M. (2008). The relationship among elementary teachers' content knowledge, attitudes, beliefs and practices. *Journal of Mathematics Teacher Education, 11*, 139-164.
- Woolfolk, A. E., & Hoy, W. K. (1990). Prospective teachers' sense of efficacy and beliefs about control. *Journal of educational Research, 82*(1), 81-91.

Appendix. Mathematics Teaching Efficacy Scale (The Final 18-Item Model)

MTPE

- 1 I will not be very effective in monitoring students' mathematics learning activities in the classroom.
나는 교실에서 학생들의 수학 학습 활동을 효율적으로 관찰하기가 힘들 것 같다.
- 2 I will not be able to teach mathematics effectively.
나는 수학을 효율적으로 가르칠 수 없을 것 같다.
- 3 I will teach elementary mathematics effectively in the future, since I well understand mathematics concepts.
나는 수학 개념을 잘 이해하고 있기 때문에, 장래에 수학을 효율적으로 가르칠 것이다.
- 4 I will have difficulty in using manipulatives to explain to students why mathematics works.
나는 학생들에게 수학이 작동하는 원리를 설명하는데 교구를 사용하는 것을 어려워 할 것 같다.
- 5 I will be able to answer students' questions about mathematics.
나는 학생들의 수학 질문에 대답을 잘 할 수 있을 것이다.
- 6 I will willingly agree to open my class to others to observe my mathematics teaching.
나는 내 수학 수업을 다른 사람들에게 공개하는데 적극적으로 동의할 것이다.
- 7 When teaching mathematics, I will like to answer students' questions.
나는 수학을 가르치면서 학생들의 질문에 대답하기를 즐겨할 것이다.
- 8 I do not know what to do to engage students to mathematics in the future.
나는 학생들을 수학에 이끌기 위하여 무엇을 해야할지 모르겠다.
- 9 I will have fear to open my mathematics class to peer teachers, staff, the principal, and parents.
나는 내 수학 수업을 동료 교사, 교장, 학부모 등에게 공개하는 것이 두려울 것 같다.
- 10 I certainly will teach mathematics well in a class to the public.
나는 공개 수업에서도 수학을 잘 가르칠 수 있을 것이다.
- 11 I will have a difficulty in adjusting mathematics lessons to the proper level for individual students.
나는 개별 학생에게 적절한 수준으로 수학 수업을 맞추는 것이 힘들 것 같다.
- 12 I will not explain some mathematical concepts very well.
나는 어떤 수학 개념은 그리 잘 설명할 수 없을 것 같다.

MTOE

- 13 When a student does better than usual in mathematics, it is because the teacher exerted extra effort.
교사가 추가적인 노력을 더하다면, 학생이 평소보다 수학을 더 잘 한다.
- 14 A student's lack of mathematical knowledge and attitudes can be overcome by good teaching.
교사의 훌륭한 가르침이 학생의 부족한 수학 지식이나 못마땅한 태도 등을 고칠 수 있다.
- 15 Teachers are the most powerful factor to student mathematics achievement.
교사는 학생의 수학 성취도에 가장 강력하게 영향을 미친다.
- 16 If a student masters a new mathematics concept quickly, this usually is because a teacher knew the necessary steps in teaching that concept.
학생이 새로운 수학 개념을 빠르게 숙달한다면, 교사가 그 개념을 가르치는데 필요한 단계를 잘 알고 있기 때문이다.
- 17 If a teacher gets students to work mathematical tasks together, then their mathematical achievement increases.
교사가 학생들로 하여금 수학 과제를 같이 하도록 하면, 학생들의 수학 성취도가 높아 진다.
- 18 A teachers' use of non-mathematical knowledge in mathematics teaching helps students understand the mathematical concept.
교사가 수학 수업에서 수학적 아닌 지식을 사용하는 것이 학생들이 수학 개념을 이해하는데 도움을 준다.

수학 교수 효능감 척도의 개발 및 정당화: 확인적 요인분석

량도형

The University of North Carolina at Greensboro

E-mail : dryang@uncg.edu

교사의 자기 효능감은 그 교사가 실제로 잘 가르칠지를 예측하는 매우 강력한 구인이다. 본 연구는 우리나라의 초등 예비교사들이 수학을 가르치는데 대한 효능감을 측정하는 정당한 척도를 개발하는 것이다. 먼저, 문헌 조사로부터 개인효능감과 결과기대감의 두 개 변인에서 58개의 문항을 개발하였고, 표본 크기 919인 자료 위에서 정규성 검사를 실시하였다. 다음에 전체 자료를 크기가 각각 419와 500인 두 개의 표본으로 나누었고, 첫째 표본에서 신뢰도 검사 및 탐색적 요인 분석을, 둘째 표본에서 확인적 요인 분석을 실시하였다. 최종적으로 개인효능감 변인에서 12 문항, 결과기대감 변인에서 6 문항 등 총 18개의 문항으로 구성된 수학교수효능감척도(MTES)를 개발하였다.

* ZDM 분류 : B59

* MSC 2000 분류 : 97C70

* Key Words : 예비교사, 요인분석, 수학, 효능감