

Psychometric Properties of the Turkish Adaptation of the Mathematics Teacher Efficacy Belief Instrument for In-Service Teachers

Bulent Cetinkaya and Ayhan Kursat Erbas

Middle East Technical University (Turkey)

Teaching efficacy beliefs have attracted researchers' attention in recent decades because of its close association with and potential impact on the implementation of new ideas and skills in education. In the present study, we have explored the psychometric properties and construct validity of the Turkish adaptation of the Mathematics Teacher Efficacy Belief Instrument developed by Enochs, Smith, & Huinker (2000) for in-service mathematics teachers. The instrument distinguishes between two dimensions of efficacy beliefs for mathematics teachers: personal mathematics teaching efficacy and mathematics teaching outcome expectancy. The sample consisted of 1355 in-service elementary school teachers and middle school mathematics teachers from 368 schools. Exploratory and confirmatory factor analysis revealed a two-factor structure similar to that found in other studies. Also, scores from the two subscales indicated acceptable internal consistency.

Keywords: self-efficacy, mathematics education, teachers, exploratory factor analysis, confirmatory factor analysis.

En décadas recientes, las creencias sobre la eficiencia de la enseñanza han atraído la atención de los investigadores debido a su cercana relación y potencial impacto en la implementación de nuevas ideas y estrategias de educación. En el presente estudio hemos explorado las propiedades psicométricas y la validez de constructo de la adaptación turca del Instrumento de Creencias de la Eficacia del Profesor de Matemáticas desarrollado por Enochs, Smith, & Huinker (2000) para profesores de matemáticas en activo. El instrumento distingue entre dos dimensiones de creencias de eficacia para profesores de matemáticas: eficacia personal en la enseñanza de las matemáticas y resultados de las expectativas en la enseñanza de las matemáticas. La muestra consistió en 1355 profesores de matemáticas de educación elemental y de escuela media en activo de 368 escuelas. Los análisis exploratorios y los análisis de factor confirmatorio revelaron una estructura de dos factores similar a la encontrada en otros estudios. Asimismo, los datos de las dos subescalas tuvieron valores de consistencia interna aceptables.

Palabras clave: auto eficacia, educación en matemáticas, profesores, análisis factorial exploratorio, análisis factorial confirmatorio.

Work reported here is based upon work supported by the Scientific and Technological Research Council of Turkey (TUBITAK) under Grant No. 107K551. Opinions expressed are those of the authors and do not necessarily represent those of TUBITAK.

Correspondence concerning this article should be addressed to Ayhan Kursat Erbas, Department of Secondary Science and Mathematics Education, Middle East Technical University, 06800 Ankara, (Turkey). Phone: +90-312-2103652. Fax: +90-312-2107971. E-mail: erbas@metu.edu.tr

The assessment of teaching efficacy beliefs has become an important concern for researchers because of its critical role on various educational outcomes, particularly on the adaptation of educational innovations (De Mesquita & Drake, 1994; Ghaith & Yaghi, 1997; Guskey & Passaro, 1994; Tschannen-Moran, Hoy, & Hoy, 1998). Research studies reveal that teachers with high teaching efficacy beliefs tend to be flexible in their teaching approaches, open to new ideas and skills, and inclined to change their teaching practices by adopting new educational ideas (Czerniak & Lumpe, 1996; Enochs & Riggs, 1990; Ghaith & Yaghi, 1997; Guskey, 1988). Although several models and instruments have been developed to measure teaching efficacy beliefs, there is a need for cross-culturally validated teaching efficacy belief scales (Brouwers, Tomic, & Stijnen, 2002; Henson, Kogan, & Vacha-Haase, 2001). The primary focus of this study was to contribute to the work on factor structure and psychometric properties of the *Mathematics Teaching Efficacy Belief Instrument* (MTEBI, Enochs, Smith, & Huinker, 2000) by translating it into Turkish and evaluating its factor structure and reliability through in-service elementary and middle school mathematics teachers in a Turkish sample. The secondary purpose of this study was to examine how the mathematics teachers' teaching experience, gender, and grade-level taught interacted with mathematics teaching efficacy beliefs. The MTEBI was selected for this study because of three reasons. The first reason is that this scale specifically focuses on mathematics teaching efficacy. Second, MTEBI and its science education version, the *Science Teaching Efficacy Belief Instrument* (STEBI, Enochs, & Riggs, 1990), have been used in various contexts with in-service and preservice teachers in different countries around the world. Although fewer studies examined MTEBI compared to STEBI, it has been translated into different languages and used in different cultural settings including Australia, South Africa, Taiwan, and Jordan. Third, only a few researchers have reported that the MTEBI has an acceptable reliability and construct validity (Alkhateeb, 2004; Cakiroglu, 2008). As a fairly new instrument, more studies are needed on the validity and reliability of the MTEBI in different populations and contexts (Enochs et al., 2000).

Teacher Efficacy and Its Assessment

Teacher efficacy is based on Bandura's (1977) social cognitive theory. He defined self-efficacy as "beliefs in one's capabilities to organize and execute the courses of action required to produce given attainments" (p. 3). The teaching efficacy belief is then conceptualized as teachers' judgment of their capacity to "influence how well students learn, even those who may be difficult or unmotivated" (Guskey & Passaro, 1994, p. 4). Bandura (1977) classifies self-efficacy in two dimensions: efficacy expectancy and outcome expectancy. Efficacy expectancy, also referred as

personal efficacy, refers to a person's beliefs about his or her capacity to successfully produce the desired outcomes in that context. Outcome expectancy, also referred as general efficacy, on the other hand, refers to a person's judgment that certain behaviors in a specific context will produce particular outcomes. The reflections of these two dimensions in relation to teacher efficacy first emerged in the studies of Ashton, Webb, and Doda (1982), and Gibson and Dembo (1984). Ashton and her colleagues' (1982) *Efficacy Vignettes* are designed to measure only personal teaching efficacy dimension of teacher efficacy. The *Teacher Efficacy Scale* (TES) developed by Gibson and Dembo (1984), on the other hand, intends to measure both dimensions of Bandura's self-efficacy theory. The first dimension refers to personal teaching efficacy and represents a teacher's belief that he or she "has the skills and abilities to bring about student learning" (p. 573). They elucidate a second dimension, general teaching efficacy, as a "belief that any teacher's ability to bring about change is significantly limited by factors external to the teacher" (p. 574). Although the TES has been criticized by some researchers concerning issues related to its factor structure, it became a starting point for developing new teacher efficacy scales (e.g., Tschannen-Moran & Hoy, 2001; Tschannen-Moran et al., 1998). One of these scales specific to a subject matter is the STEBI (Enochs & Riggs, 1990). Based on the TES, Riggs and Enochs (1989) constructed the STEBI, a new instrument specific to science teachers, and then they modified it and developed the MTEBI (Enochs et al., 2000). Contribution of these instruments to the teaching self-efficacy research was significant not only because they address the weaknesses of the TES, but also they were solely designed for a specific subject matter (Liu, Jack, & Chiu, 2007).

Psychometric Properties of the MTEBI and the STEBI

Psychometric properties of the MTEBI and the STEBI are discussed together here since (a) the constructs are similar, indeed most of the time the only change made was to replace the term 'science' with the term 'mathematics', and (b) as a new instrument, available data on psychometric properties of the MTEBI are limited. The STEBI instrument originally consisted of 25 items. After analyzing its factor structure and reliability, Enochs and Riggs (1990) dropped two items because of their cross-loadings. Enochs et al. (2000) subsequently conducted an item analysis for the revised version of the 23-item STEBI and developed a 21-item MTEBI by deleting two low correlated items.

The MTEBI is comprised of two subscales, namely personal mathematics teaching efficacy (PMTE, 13 items) and mathematics teaching outcome expectancy (MTOE, 9 items). Using a sample of 324 elementary preservice mathematics teachers, Enochs et al. (2000) reported internal consistencies (Cronbach's alpha) for PMTE and MTOE

scales as .88 and .77, respectively. They also reported the independence of the two scales through confirmatory factor analysis. Studies utilizing MTEBI in the United States generally cite alpha values reported by Enochs et al. (2000) without examining the reliability of their own MTEBI data (e.g., Gresham, 2008; Swars, Daane, & Giesen, 2006). However, several other researchers found high reliabilities for the two subscales of teaching efficacy belief for mathematics and science, ranging from .77 to .92 for PTE, and from .65 to .76 for TOE. Table 1 shows the internal consistency of the MTEBI and the STEBI reported in some of these studies.

Enochs and Riggs (1990) and Enochs et al. (2000) conceptualized science and mathematics teaching efficacy beliefs in two dimensions, namely personal teaching efficacy and teaching outcome expectancy. This two-factor structure for the STEBI for use in the US and some other countries has been established in science education literature (e.g., Bleicher, 2004; Mji & Kiviet, 2003; Mulholland, Dorman, & Odgers, 2004; Tekkaya, Cakiroglu, & Ozkan, 2004). In some of these studies, however, researchers offered minor changes. For example, Bleicher (2004) reported that removing the word “some” improved the item loadings and item total correlations for the items 10 and 13. A comprehensive review of research revealed that although the MTEBI has been used extensively in the studies, its construct validity was explored in only one study (Alkhateeb, 2004). He administered the Arabic translation of the 21-

item MTEBI to 144 undergraduate students in a school of education in Jordan. He found that two factors existed corresponding to two original dimensions that accounted for 41% of the total variance, and all the items loaded on the factors were as expected.

The review of literature reveals that although the MTEBI and the STEBI have been used widely, majority of the studies were conducted with pre-service teachers and only a few with in-service teachers. Also, only a few study examined psychometric properties of the MTEBI. Therefore, there is a strong need to examine the reliability of the construct of mathematics teaching self-efficacy and to extend the validity of the MTEBI to in-service teachers.

Method

Participants

The participants of the study were 1355 in-service elementary school teachers and middle school mathematics teachers from 368 schools throughout Turkey. While 1098 (81%) of the participants were teaching in public schools, 257 (19%) of them were teaching in private schools. The average age of the participants ranged from 21 to 67 ($M = 37.4$, $SD = 9.3$). About 65% of the participants were less than 40 years old. Table 2 shows some demographic characteristics of the sample.

Table 1
Cronbach's alpha reliability coefficients reported in some of the previous studies

Study	Participants	Inventory	Items	Subscales	
				PTE	TOE
Alkhateeb (2004)	144 Jordanian undergraduate students	MTEBI	21 Items, original scale	.84	.75
Cakiroglu (2008)	245 elementary preservice teachers in US and Turkey	MTEBI	21 Items, original scale	.77	.65
Enochs et al. (2000)	324 elementary preservice teachers	MTEBI	21 Items, original scale	.88	.77
Enochs & Riggs (1990)	212 American preservice teachers	STEBI	23 Items, original scale	.90	.76
Bleicher (2004)	290 American elementary preservice science teachers	STEBI	23 Items, original scale	.87	.72
Tekkaya, et al. (2004)	299 Turkish elementary preservice science teachers in US and Turkey	STEBI	23 Items, original scale	.84	.76
Mji & Kiviet (2003)	200 South African elementary teachers	STEBI	25 Items, all the items of the first version	.92	.73
Mulholland, et al. (2004)	314 Australian elementary preservice teachers	STEBI	21 Items, 2 deleted items were not reported	.83	.74
Liu, et al. (2007)	282 Taiwanese elementary science teachers	STEBI	16 Items, deleted items*: 7, 10, 11, 14, 17, 21, 22, 20, 25	.82	.81

* These items deleted because of low item-total correlations ($< .32$).

Table 2
Characteristics of the participants

Characteristics	<i>f</i> (%)
<i>Level teaching</i>	
Elementary school (grades 1-5)	827 (61)
Middle school (grades 6-8)	528 (39)
<i>Gender</i>	
Female	749 (55.3)
Male	606 (44.7)
<i>Experience in teaching (years)*</i>	
0-2	139 (10.3)
3-5	185 (13.7)
6-9	270 (19.9)
10-19	459 (33.9)
20 and above	298 (22)

* Data for 4 participants was missing.

Instrument

In this study, in-service elementary and middle school teachers' mathematics teaching self-efficacy beliefs were measured using an extended and translated (into Turkish) version of the Mathematics Teaching Efficacy Belief Instrument (MTEBI) for pre-service teachers (Enochs et al., 2000). MTEBI for pre-service teachers was adopted from the Science Teaching Efficacy Belief Instrument (STEBI-B) for pre-service teachers (Enochs & Riggs, 1990; Riggs & Enoch, 1989). The MTEBI for pre-service teachers is a 21-item self-report scale developed to measure pre-service teachers' mathematics teaching efficacy beliefs and their outcome expectancy. Each item is rated on a 5-point Likert type scale ranging from 5 (strongly agree) to 1 (strongly disagree). The MTEBI for pre-service teachers consisted of two subscales: the personal mathematics teaching efficacy beliefs (PMTE) (13 items, e.g., I wonder if I have the necessary skills to teach mathematics) and the mathematics teaching outcome expectancy (MTOE) (8 items, e.g., When a low achieving child progresses in mathematics, it is usually due to extra attention given by the teacher). Possible scores on the PMTE scale range from 13 to 65 and MTOE scores may range from 8 to 40. The higher the score on the PMTE scale, the stronger the personal beliefs in one's efficacy as a mathematics teacher. Similarly, the higher the score on the MTOE scale, the higher the expectations of the outcomes of mathematics teaching.

For this study, we decided to include two outcome expectancy items used by Enoch and Riggs (1990) but dropped later by Enoch et al. (2000) because of low item-total item correlation. We also included another two outcome expectancy items mentioned in Riggs and Enoch (1989) but not covered by Enoch et al. (2000). Thus, in the final form, the MTEBI for in-service teachers consisted of 25

items. The back-translation design (Hambleton, 2005) guided the adaptation of the MTEBI into Turkish. First, a bilingual mathematics education professor translated the original items into Turkish. During this stage, Turkish translations and adaptations of the MTEBI for pre-service teachers (Cakiroglu, 2008) and the STEBI for in-service elementary teachers (Tekkaya et al., 2004) were very useful. Special attention was paid to semantic, idiomatic, and conceptual equivalence to preserve overall meaning and nuances and to ensure cultural and psychological equivalence. Since our purpose was to measure in-service teachers' teaching efficacy rather than that of pre-service teachers, use of concepts, words, and expressions that would make sense for practicing teachers were ascertained. For example, present rather than future tense (as it is the case with the prospective teachers) was used. Furthermore, negatively worded items were translated as such. The items were then back-translated into English by another bilingual mathematics education professor. Two translators compared the original and the back-translated versions for any inconsistency in meaning and then adjustments to the Turkish version were made accordingly.

Data Analysis

In the data analysis, we first examined the reliability of teacher responses to individual items and to the subscales suggested in the previous studies (e.g., Enoch et al., 2000; Enoch & Riggs, 1990; Riggs & Enoch, 1989) according to item-total correlations and alpha coefficients, respectively. Initially, the data were examined for missing values and normality. It was decided to exclude missing cases listwise. Except for a few cases, the missing data were random and deleting them did not lower the sample size significantly. Moreover, it was decided not to carry out any type of data transformation to improve the normality although some of the item scores in the MTEBI for in-service teachers were skewed. For example, the logarithmic transformation created significant skewness in some of the items while there was no problem of skewness before the transformation applied.

After the listwise deletion of missing cases, the remaining sample ($N = 1119$) was randomly divided into two subsamples. Both subsamples were matched based on the level taught, gender and teaching experience. Data from the first subsample ($n = 552$) were subjected to an exploratory factor analysis (EFA) using SPSS 17. EFA was performed using the principal component analysis (PCA). Various methods have to be considered in deciding on the number of factors to be retained after conducting the principal components analysis. These methods are the parallel analysis (Horn, 1965), the minimum average partial method (MAP) (Velicer, 1976), Kaiser-Guttman criterion (i.e., eigenvalues ≥ 1), scree test, and theoretical interpretability of factors (Field, 2005; Tabachnick & Fidell, 2007). SPSS syntaxes provided by O'Connor (2000) were used to perform the MAP test and parallel analysis. Data

from the second subsample ($n = 567$), on the other hand, were used to corroborate the identified factor structure through confirmatory factor analysis (CFA) using AMOS 16 (Arbuckle, 2007). As the chi-square statistic is extremely sensitive to sample size, CFI, TLI, RMSEA, and SRMR fit indices of the hypothesized latent factor structure for the observed data were examined while evaluating the model fit (Byrne, 2001; Hu & Bentler, 1999; Kline, 2005). Finally, the internal consistency of the resulting instrument was examined with a one-way between-subject multivariate analysis of variance (MANOVA) conducted using the whole sample to see whether the teachers differ on the identified factors according to the level they taught, their gender and teaching experience.

Results

Table 3 summarizes the results of the reliability analysis of the MTEBI for in-service teachers based on the two-factor model suggested in the previous studies (e.g., Enochs et al., 2000; Enochs & Riggs, 1990; Riggs & Enochs, 1989). The computed alpha values suggested that scores from the first scale produced a good and acceptable reliability coefficient while this was low for the second scale scores. This result might have been caused by four items (10, 13, 20, and 25) in the second scale because they had rather low item-total correlations. Thus, these four items were deleted and excluded from further analyses.

Table 3
Summary of reliability estimates for MTEBI for in-service teachers

Subscale/Item	Item-total correlations
Factor 1 - Personal mathematics teaching efficacy (PMTE) ($\alpha = .82$)	
2. I am continually finding better ways to teach mathematics.	.45
*3. Even when I try very hard, I don't teach mathematics as well as I do most subjects.	.42
5. I know how to teach mathematics concepts effectively.	.46
*6. I am not very effective in monitoring mathematics activities.	.44
*8. I generally teach mathematics ineffectively.	.53
12. I understand mathematics concepts well enough to be effective in teaching elementary mathematics.	.53
*17. I find it difficult to use manipulatives to explain to students why mathematics works.	.47
18. I am typically able to answer students' mathematics questions.	.47
*19. I wonder if I have the necessary skills to teach mathematics.	.61
*21. Given a choice, I would not invite the principal to evaluate my mathematics teaching.	.24
*22. When a student has difficulty understanding a mathematics concept, I am usually at a loss as to how to help the student understand it better.	.59
23. When teaching mathematics, I usually welcome student questions.	.45
*24. I don't know what to do to turn students on to mathematics.	.46
Factor 2 - Mathematics teaching outcome expectancy (MTOE) ($\alpha = .63$)	
1. When a student does better than usual in mathematics, it is often because the teacher exerted a little extra effort.	.29
4. When the mathematics grades of students improve, it is most often due to their teacher having found a more effective teaching approach.	.41
7. If students are underachieving in mathematics, it is most likely due to ineffective mathematics teaching.	.26
9. The inadequacy of a student's mathematics background can be overcome by good teaching.	.29
*10. The low mathematics achievement of some students cannot generally be blamed on their teachers.	.04
11. When a low achieving child progresses in mathematics, it is usually due to extra attention given by the teacher.	.43
*13. Increased effort in mathematics teaching produces little change in some students' mathematics achievement.	.18
14. The teacher is generally responsible for the achievement of students in mathematics.	.34
15. Students' achievement in mathematics is directly related to their teacher's effectiveness in mathematics teaching.	.42
16. If parents comment that their child is showing more interest in mathematics at school, it is probably due to the performance of the child's teacher.	.34
*20. Effectiveness in mathematics teaching has little influence on the achievement of students with low motivation.	.19
*25. Even teachers with good mathematics teaching abilities cannot help some kids learn mathematics.	.22

Note. Items marked with an "*" are negatively worded and need to be reversed in scoring.

Principal Component Analysis (PCA)

None of the correlation coefficients between each pair of items in R-matrix was particularly large; therefore, there was no need to consider eliminating any items at this stage. Also, the determinant of the R-matrix was .012, indicating that multicollinearity was not a problem for the data set. Furthermore, the Kaiser-Meyer-Olkin test showed that the correlation matrices were factorable (KMO = .85) and the quality of the sampling was good. Bartlett's test of sphericity was highly significant ($\chi^2(210) = 2384.87, p < .001$). Then, PCA was performed on the 21 items from the MTEBI for in-service teachers. The initial analysis extracted 5 factors with eigenvalues greater than one. This solution accounted for 50.32% of the total variance. The eigenvalues of the first five factors were: 4.66, 2.24, 1.41, 1.18, and 1.07. Next, the data were analyzed by orthogonal (varimax) and oblique (direct oblimin) methods of transformation. Both transformations revealed four factors similar to the initial analysis.

We observed that Kaiser-Guttman criterion overestimated the number of factors to be retained. For example, most of the items loaded on the first and second factors accounted for 22.2% and 10.67% of the total variance, respectively. Also, several items loaded on more than one factor according to the Kaiser-Guttman criterion. Additionally, considering the fact that resulting communalities (after extraction) were either equal to or all less than .62 with an average of .50, we decided not to rely on the Kaiser-Guttman criterion as suggested by Field (2005). On the other hand, the scree plot revealed three or four factors.

In deciding the number of factors to retain, rather than more rule-based traditional approaches we reported above whereby data distributions may have affected factors and eigenvalues, we preferred to rely on the MAP test and parallel analysis because these statistically based methods are considered to produce optimal solutions for determining the dimensionality of a construct (Glorfeld, 1995; Henson & Roberts, 2006; O'Connor, 2000; Zwick & Velicer, 1986). In the parallel analysis, five thousands random datasets were created, each of which had 552 cases and 21 variables. In 95% of the datasets generated, the first four eigenvalues were equal or less than 1.37, 1.3, 1.25, and 1.21. Thus, the parallel analysis suggested that three factors underlie the measure of efficacy. On the other hand, the MAP revealed average squared partial correlations (ASPC) of .042 with no components extracted; .015 with one component extracted; .012 with two components extracted; and, .013 with three components extracted. Contrary to the parallel analysis, a two-factor solution was suggested by the MAP as the smallest ASPC was associated with the second component. We decided that the two-factor solution suggested by the MAP would be more meaningful and theoretically interpretable considering the fact that the original instruments (i.e., the STEBI and MTEBI for in-service and pre-service teachers) have been created based

on two theoretical constructs after Bandura's (1977) two dimensions of teacher efficacy (i.e., outcome expectations and self-efficacy expectations) and previous research conducted using these instruments (e.g., Bleicher, 2004; Cakiroglu, 2008; Enochs & Riggs, 1990; Enochs et al., 2000; Riggs, & Enochs, 1989) has confirmed these two lower order factors extracted through exploratory and confirmatory factor analyses.

Although Enochs et al. (2000) reported the independence of the two scales through confirmatory factor analysis, Enochs and Riggs (1990), and Bleicher (2004) reported a modest correlation between the two factors. The PCA for a two-factor solution was conducted using oblique rotation (direct oblimin with delta = 0) with Kaiser normalization. The results are presented in Table 4, wherein factor loadings of pattern and structure matrices and communalities are shown. The two factors accounted for 32.87% of the total variance, with eigenvalues of 4.66 and 2.24 for factors 1 and 2, respectively. These factors were moderately correlated ($r = .22$ at a delta value of 0), signifying that they were related but independent constructs. As seen in Table 4, the two-factor solution revealed a simple structure of the MTEBI for in-service teachers that is similar to that reported by Enochs et al. (2000). Based on this, Factor 1 was named as *personal mathematics teaching efficacy* (PMTE) and Factor 2 as *mathematics teaching outcome expectancy* (MTOE) (Enochs et al., 2000; Enochs & Riggs, 1990; Riggs & Enochs, 1989).

It has been suggested that variables with pattern coefficients of .32 or larger are generally acceptable for item inclusion (Tabachnick & Fidell, 2007). Based on this suggestion, Item 21 was removed because its pattern coefficient was less than .32. This means that Factor 1 (i.e., PMTE) was made up of 12 items, namely Item 2, Item 3, Item 5, Item 6, Item 8, Item 12, Item 17, Item 18, Item 19, Item 22, Item 23, and Item 24. On the other hand, Factor 2 (i.e., MTOE) had 8 items, namely Item 1, Item 4, Item 7, Item 9, Item 11, Item 14, Item 15, and Item 16. It may also be observed in Table 4 that Item 2 (PMTE = .50 and MTOE = .32), Item 9 (PMTE = .40 and MTOE = .33) and Item 4 (PMTE = .34 and MTOE = .64) had cross factor loadings. These cross loadings can be neglected, since the primary loadings were significantly higher than the secondary ones. Based on the analysis reported here, subsequent computations involved the 20-item MTEBI for in-service teachers.

Confirmatory Factor Analysis

Confirmatory factor analysis (CFA) was conducted to examine the construct validity of the self-efficacy scale. The CFA model indicated a poor level fit to the given data in terms of chi-square, CFI, and TLI indices $\chi^2(169 df, N = 567) = 600.22, p < .001, CFI = .84, TLI = .82$. This was in spite of the RMSEA (.067 with 90% CI .061 – .073) and SRMR (.067) indices which fell within an acceptable range (Hu & Bentler, 1999). Modification indices suggested

Table 4

Factor loadings (from pattern and structure matrices) and communalities (η^2) of the items in MTEBI for in-service teachers for the principal component analysis after direct oblimin rotation

Item#	Pattern Matrix		Structure Matrix		η^2	Item#	Pattern Matrix		Structure Matrix		η^2
	F1	F2	F1	F2			F1	F2	F1	F2	
22	-.71	.16	-.68	.01	.48	11	.05	.68	.20	.69	.48
19	-.67	.00	-.67	-.15	.45	16	-.04	.66	.10	.65	.43
8	-.59	-.04	-.60	-.17	.36	15	-.01	.61	.12	.61	.37
6	-.58	.10	-.55	-.03	.32	4	.22	.60	.34	.64	.46
18	.57	.08	.59	.21	.35	14	-.13	.59	.00	.56	.33
5	.56	.06	.57	.18	.33	1	.11	.51	.22	.53	.29
12	.55	.17	.59	.29	.37	7	-.05	.40	.04	.39	.16
17	-.54	.11	-.52	.00	.28						
23	.54	.00	.54	.12	.29						
3	-.53	.02	-.53	-.09	.28						
24	-.52	.09	-.50	-.02	.25						
2	.45	.23	.50	.32	.3						
9	.34	.26	.40	.33	.22						
21	-.30	-.04	-.31	-.10	.1						

Note. Salient loadings are $\geq .3$. Numbers in **bold** represent the highest salient factor loadings on a factor. Numbers in *italics* specify the second highest salient loadings on a factor. Factor labels: Factor I, personal mathematics teaching efficacy (PMTE); Factor II, Mathematics teaching outcome expectancy (MTOE).

that error variances of several items could be correlated to increase model fit. Based on the modification indices and theoretical relevance, links between the items 6 and 8, 12 and 18, 14 and 15, and 18 and 23 were allowed. The error covariance between the first three pairs of the items was likely caused by content overlap. For example, contents of the items 14 and 15 focused directly on teachers' performance as a cause of students' progress in mathematics; and the contents of the items 12 and 18 have to do with teachers' beliefs about their mathematical knowledge. The error covariance between Items 6 and 8 might be because of the close semantic likeness between these items in the Turkish language. A careful consideration of the contents of Item 18 (I am typically able to answer students' mathematics questions) and Item 23 (When teaching mathematics, I usually welcome student questions) suggested that the error covariance between them might be the bias caused by the social desirability of the items. After modification, the values of CFI (.90), TLI (.88), SRMR (.06) and RMSEA (.056 with 90% CI .05 – .062) fit indices indicated the tested model had an acceptable fit to the data. The chi-square statistic $\chi^2(165 \text{ df}, N = 567) = 455.49, p < .001$ of the modified CFA model was substantially lower than that of the original model.

The results of factor loadings and measurement error variances of the modified CFA model are provided in Figure 1. All indicators in the model had statistically significant unstandardized factor loadings to their common latent factors ($p < .001$), corroborating the presence of significant

relationships among measured indicators and their latent variables. Also, except for Item 7 in MTOE, all indicators had satisfactory standardized factor loadings on their common latent factor. Item 7 was not deleted because its factor loading .29 was not substantially low. Bivariate correlation between MTOE and MTOE was statistically significant ($r = .44$).

Internal Consistency of the MTEBI for In-service Teachers

Reversing the negatively worded items, the internal consistency of scores from the overall scale (Cronbach's alpha) was found to be .82 ($n = 1128$). Moreover, Cronbach's alpha coefficients for the factors personal mathematics teaching efficacy (PMTE) and mathematics teaching outcome expectancy (MTOE) were .83 ($n = 1158$) and .70 ($n = 1298$), respectively. Item total correlations of all items with the rest of the items ranged from .42 to .61 in PMTE and from .30 to .50 in MTOE.

Further Validation: Background Variable Associations with MTEBI for In-service Teachers' Scores

Barlett's test of sphericity ($\chi^2 = 237.07, df = 2, p < .001$) indicated that MANOVA was warranted. Also, the Levene's test suggested heterogeneity of variances for both personal mathematics teaching efficacy (PMTE) and mathematics teaching outcome expectancy (MTOE) variables. Thus, a

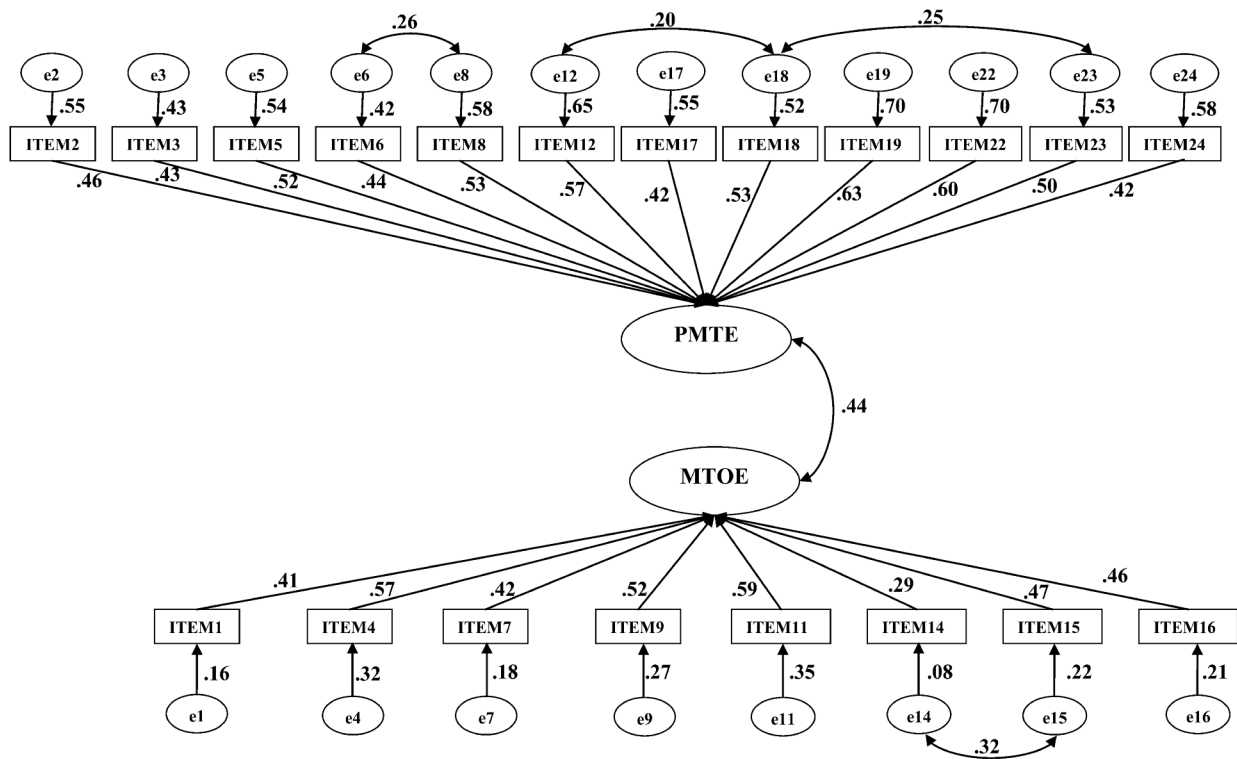


Figure 1. Standardized loadings for the two-factor model of mathematics teaching self-efficacy beliefs scale (All coefficients are significant at $p < .01$).

one-way between-subject multivariate analysis of variance (MANOVA) was conducted on two dependent variables: PMTE and MTOE. The independent variables were the grade-level taught (elementary school teachers; mathematics teachers teaching grades 6 to 8); gender and (years of) experience in teaching (0-2, 3-5, 6-9, 10-19, and 20 and above). A statistically significant Box’s M test ($p = .012$) suggested unequal variance-covariance matrices of the dependent variables across levels of experience in teaching, gender groups, and the grade-level taught and thus compelled the use of Pillai’s trace in assessing the multivariate effect.

With the use of Pillai’s trace, the combined DVs were significantly affected only by teacher’s gender; Pillai’s trace = .014, $F(2,1104) = 8.028$, $p < .001$, partial $\eta^2 = .014$. The level taught and experience in teaching were not significant; Pillai’s trace $< .001$, $F(2,1104) = 0.115$, $p = .891$, partial $\eta^2 < .001$ and Pillai’s trace = .004, $F(8,2210) = 0.546$, $p = .822$, partial $\eta^2 = .002$. However, no interaction between the independent variables was statistically significant. When the results for the dependent variables were considered separately, the only difference to reach statistical significance, using a Bonferroni adjusted alpha level of .025, was personal mathematics teaching efficacy (PMTE): $F(1,1105) = 9.42$, $p = .002$. An inspection of mean scores suggested that female teachers reported slightly higher levels of higher personal mathematics teaching efficacy ($M = 48$, $SD = 6.28$) than males ($M = 47$, $SD = 6.62$).

Discussion

Although teaching efficacy plays an important role in effective mathematics teaching, its measurement is still being questioned because of validity and reliability issues. This study therefore explored the psychometric properties and construct validity of the Turkish translation of the *Mathematics Teacher Efficacy Belief Instrument* (MTEBI) developed by Enochs et al. (2000) for in-service mathematics teachers with added four items used in previous studies. An initial reliability analysis based on the two-factor model suggested in previous studies (e.g., Enochs et al., 2000; Enochs & Riggs, 1990; Riggs & Enochs, 1989) suggested the deletion of four items (10, 13, 20, and 25) from MTOE because of low item-total correlations, leaving 21 items for further analysis. Deletion of these items was consistent with Riggs and Enochs (1989) and Enochs and Riggs (1990).

Contrary to the Horn’s parallel analysis suggesting that three factors underlies the measure of efficacy; a two-factor solution was suggested by Velicer’s MAP test. A two-factor solution was tested by PCA considering that it would be theoretically more relevant as the two factors empirically mirrored two self-efficacy dimensions for (mathematics) teachers: personal mathematics teaching efficacy (PMTE) and mathematics teaching outcome expectancy (MTOE). Although it was not consistent with previous studies, results of EFA suggested the deletion of Item 21. Considering the

fact that all previous studies reviewed have been conducted with preservice teachers except for one by Mji and Kiviet (2003) (see Table 1), inclusion of this item can be problematic for in-service teachers and needs further validation, particularly in smaller sample sizes. Thus, the 20-item MTEBI for in-service mathematics teachers was found to measure two dimensions of efficacy beliefs of in-service mathematics teachers.

The confirmatory factor analysis suggested that the two-factor model showed acceptable levels of model fit similar to those reported by Enochs et al. (2000). Therefore, it was concluded that the items in the Turkish version of MTEBI for in-service teachers measure two latent dimensions: PMTE and MTOE. Furthermore, the factor structure provides evidence for the structural aspects of construct validity, since the scores from the Turkish version of the MTEBI were consistent with Bandura's (1977) two dimensions of teacher self-efficacy, namely outcome expectations and self-efficacy expectations.

On the other hand, the internal consistency reliabilities of PMTE and MTOE were found to be very good and good, respectively, similar to those reported in other studies (Alkhateeb, 2004; Bleicher, 2004; Henson et al., 2001, Mji & Kiviet, 2003; Mulholland et al., 2004; Tekkaya et al., 2004). The low alpha value score for the MTOE dimension was expected since it has been criticized by several researchers that it would not be an appropriate construct to measure teacher efficacy (e.g., Guskey & Passaro, 1994; Henson et al., 2001).

An examination of whether male and female teachers differed in terms of the two factors of the MTEBI for in-service teachers revealed that a statistically significant difference existed in teachers' personal teaching efficacy beliefs in favor of females. Yet, there is no agreement among the findings of the studies in the literature about the gender differences on personal teaching efficacy beliefs. There are several studies reporting no significant difference between males and females with respect to their personal teaching efficacy for teaching science and mathematics (e.g., Cakiroglu, 2008; Cakiroglu, Cakiroglu, & Boone, 2005; Guskey & Passaro, 1994; Mulholland et al., 2004). However, there are some revealing significant differences with respect to females (e.g., Anderson, Greene, & Lowen, 1988; Evans & Tribble, 1986) and others with respect to males (e.g., Bleicher, 2004; Enochs & Riggs, 1990; Riggs, 1991). On the other hand, the results revealed no statistically significant gender effect in teachers' teaching outcome expectancy beliefs. This finding is consistent with the results of other studies (e.g., Bleicher, 2004; Cakiroglu, 2008; Cakiroglu et al., 2005; Guskey & Passaro, 1994; Mulholland et al., 2004). The results also showed that teaching experience and the level taught (i.e., elementary vs. middle grades) had no significant effect on both personal mathematics teaching efficacy and mathematics teaching outcome expectancy. This finding is consistent with the findings of

other studies (Anderson et al., 1988; Bleicher, 2004; Enochs & Riggs, 1990; Guskey & Passaro, 1994).

To sum up, this study was an attempt to contribute to the international work on the evaluation of psychometric properties of the MTEBI and its science education version (STEBI). In general, the study supported the use of the MTEBI as a scale to measure mathematics teaching efficacy belief in a Turkish population like in other cultures and populations. Similar to the results of the studies in Western and non-Western populations, the translated Turkish version of the MTEBI for in-service teachers possessed adequate psychometric properties and construct validity for providing precise and valid information about efficacy beliefs of in-service (elementary and middle school) mathematics teachers. Nevertheless, the mathematics teaching outcome expectancy (MTOE) dimension needs further empirical validation as also suggested by Alkhateeb (2004) and Henson et al. (2001).

References

- 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. doi:10.2466/pr0.94.3.833-838
- Anderson, R. N., Greene, M. L., & Loewen, P. S. (1988). Relationships among teachers' and students' thinking skills, sense of efficacy, and student achievement. *The Alberta Journal of Educational Research, 34*, 148-165.
- Arbuckle, J. L. (2007). Amos (Version 16) [Computer software]. Spring House, PA: Amos Development Corporation.
- Ashton, P., Webb, R., & Doda, C. (1982). *A study of teachers' sense of efficacy: Final report, Vols. 1 & 2*. Gainesville, FL: University of Florida.
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*(2), 191-215. doi:10.1037//0033-295X.84.2.191
- Bleicher, R. E. (2004). Revisiting the STEBI-B: Measuring self-efficacy in preservice elementary teachers. *School Science and Mathematics, 104*(8), 383-391. doi:10.1111/j.1949-8594.2004.tb18004.x
- Brouwers, A., Tomic, W., & Stijnen, S. (2002). A confirmatory factor analysis of scores on the teacher efficacy scale. *Swiss Journal of Psychology, 61*(4), 211-219. doi:10.1024//1421-0185.61.4.211
- Byrne, B. M. (2001). *Structural equation modeling with Amos: Basic concepts, applications, and programming*. Mahwah, NJ: Erlbaum.
- 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. doi:10.1080/02607470701773457
- Cakiroglu, J., Cakiroglu, E., & Boone, W. J. (2005). Pre-service teacher self-efficacy beliefs regarding science teaching: A comparison of pre-service teachers in Turkey and the USA. *Science Educator, 14*, 31-40.

- Czerniak, C. M., & Lumpe, A. T. (1996). Relationship between teacher beliefs and science education reform. *Journal of Science Teacher Education*, 7(4), 247-266. doi:10.1007/BF00058659
- De Mesquita, P. B., & Drake, J. C. (1994). Educational reform and self-efficacy beliefs of teachers implementing nongraded primary school programs. *Teaching and Teacher Education*, 10(3), 291-302. doi:10.1016/0742-051X(95)97311-9
- Enochs, L. G., & Riggs, I. M. (1990). Further development of an elementary science teaching efficacy belief instrument: A preservice elementary scale. *School Science and Mathematics*, 90, 695-706. doi:10.1111/j.1949-8594.1990.tb12048.x
- Enochs, L. G., Smith, P. L., & Huinker, D. (2000). Establishing factorial validity of the Mathematics Teaching Efficacy Beliefs Instrument. *School Science and Mathematics*, 100(4), 194-202. doi:10.1111/j.1949-8594.2000.tb17256.x
- Evans, E. D., & Tribble, M. (1986). Perceived teaching problems, self-efficacy, and commitment to teaching among preservice teachers. *Journal of Educational Research*, 80(2), 81-85.
- Field, A. (2005). *Discovering statistics using SPSS*. London, UK: Sage.
- Ghaith, G., & Yaghi, M. (1997). Relationships among experience, teacher efficacy and attitudes toward the implementation of instructional innovation. *Teaching and Teacher Education*, 13(4), 451-458. doi: 10.1037/0022-0663.76.4.569
- Gibson, S., & Dembo, M. H. (1984). Teacher efficacy: A construct validation. *Journal of Educational Psychology*, 76(4), 569-582. doi:10.1037/0022-0663.76.4.569
- Glorfeld, L. W. (1995). An improvement on Horn's parallel analysis methodology for selecting the correct number of factors to retain. *Educational and Psychological Measurement*, 55(3), 377-393. doi:10.1177/0013164495055003002
- Gresham, G. (2008). Mathematics anxiety and mathematics teacher efficacy in elementary pre-service teachers. *Teaching Education*, 19(3), 171-184. doi:10.1080/10476210802250133
- Guskey, T. R. (1988). Teacher efficacy, self-concept, and attitudes toward the implementation of instructional innovation. *Teaching and Teacher Education*, 4(1), 63-69. doi:10.1016/0742-051X(88)90025-X
- Guskey, T. R., & Passaro, P. D. (1994). Teacher efficacy: A study of construct dimensions. *American Educational Research Journal*, 31(3), 627-643. doi:10.3102/00028312031003627
- Hambleton, R. K. (2005). Issues, designs, and technical guidelines for adapting tests into multiple languages and cultures. In R. K. Hambleton, P. F. Merenda, & C. D. Spielberger (Eds.), *Adapting educational and psychological tests for cross-cultural assessment* (pp. 3-38). Mahwah, NJ: Lawrence Erlbaum.
- Henson, R. K., & Roberts, J. K. (2006). Use of exploratory factor analysis in published research: Common errors and some comment on improved practice. *Educational and Psychological Measurement*, 66(3), 393-416. doi:10.1177/0013164405282485
- Henson, R. K., Kogan, L. R., & Vacha-Haase, T. (2001). A reliability generalization study of the teacher efficacy scale and related instruments. *Educational and Psychological Measurement*, 61(3), 404-420. doi:10.1177/00131640121971284
- Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, 30(2), 179-185. doi:10.1007/BF02289447
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1-55. doi:10.1080/10705519909540118
- Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York, NY: Guilford.
- Liu, C. J., Jack, B. M., & Chiu, H. L. (2007). Taiwan elementary teachers' views of science teaching self-efficacy and outcome expectations. *International Journal of Science and Mathematics Education*, 6(1), 19-35. doi:10.1007/s10763-006-9065-4
- Mji, A., & Kiviet, A. M. (2003). Psychometric characteristics of the Science Teaching Efficacy Belief Inventory in South Africa. *Psychological Reports*, 92, 325-332. doi:10.2466/PRO.92.1.325-332
- Mulholland, J., Dorman, J. P., & Odgers, B. M. (2004). Assessment of science teaching efficacy of preservice teachers in an Australian university. *Journal of Science Teacher Education*, 15(4), 313-331. doi: 10.1023/B:JSTE.0000048334.44537.86
- O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods, Instrumentation, and Computers*, 32, 396-402. doi:10.3758/BF03200807
- Riggs, I. M. (1991, April). *Gender differences in elementary science teacher self efficacy*. Paper presented at the annual meeting of the American Educational Research Association. Chicago, IL.
- Riggs, I. M., & Enochs, L. G. (1989, March-April). *Toward the development of an elementary teacher's science teaching efficacy belief instrument*. Paper presented at the Annual Meeting of the National Association for Research in Science Teaching (62nd). San Francisco, CA.
- 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. doi:10.1111/j.1949-8594.2006.tb17921.x
- Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston, MA: Pearson Education.
- Tekkaya, C., Cakiroglu, J., & Ozkan, O. (2004). Turkish preservice science teachers' understanding of science and their confidence in teaching it. *Journal of Education for Teaching: International Research and Pedagogy*, 30(1), 57-66. doi:10.1080/0260747032000162316
- Tschannen-Moran, M., & Hoy, A. W. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and Teacher Education*, 17(7), 783-805. doi:10.1016/S0742-051X(01)00036-1
- Tschannen-Moran, M., Hoy, A. W., & Hoy, W. K. (1998). Teacher efficacy: Its meaning and measure. *Review of Educational Research*, 68(2), 202-248. doi:10.3102/00346543068002202
- Velicer, W. F. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika*, 41(3), 321-327. doi:10.1007/BF02293557

Zwick, W. R., & Velicer, W. F. (1986). Comparison of the rules for determining the number of components to retain. *Psychological Bulletin*, *99*(3), 432-442. doi:10.1037//0033-2909.99.3.432

Received April 5, 2010
Revision received October 8, 2010
Accepted November 22, 2010