Development and Validation of a Scale for Measuring Mathematics Teaching Self-Efficacy for Teachers in the Sultanate of Oman

Hussain Alkharusi
Dr., correspondence author, Sultan Qaboos University, Sultanate of Oman, Oman, *hussein5@squ.edu.om*

Said Aldhafri
Sultan Qaboos University, Sultanate of Oman, Oman

Khoula Al-Hosni
Ministry of Education, Sultanate, Oman

Saleh Al-Busaidi
Sultan Qaboos University, Sultanate of Oman, Oman

Bader Al-Kharusi
Sultan Qaboos University, Sultanate of Oman, Oman

Abdullah Ambusaidi
Sultan Qaboos University, Sultanate of Oman, Oman

Marwa Alrajhi
Sultan Qaboos University, Sultanate of Oman, Oman

A scale for measuring self-efficacy for teaching mathematics in grades 5 to 10 was developed in this study for teachers in Oman. The participants were 328 mathematics teachers randomly selected from five educational governorates in the Sultanate of Oman. Factorial structure of the scale revealed three subscales: self-efficacy for understanding the mathematics content, self-efficacy for teaching the mathematics content, and general teaching self-efficacy. The three subscales showed strong internal consistency and sufficient evidence of construct validity and concurrent validity. The scale has potential uses for both educational and research purposes.

Keywords: self-efficacy, teaching efficacy, mathematics efficacy, validity, reliability, scale development

INTRODUCTION

Self-efficacy is concerned with individuals' beliefs in their capabilities to conduct courses of action required to achieve given outcomes (Bandura, 2006, p. 1). It influences the choice of activities individuals make, the effort they invest, and how long they will persist in stressful situations (Pajares, 1997). A strong sense of efficacy can be motivating and persevering while a weak sense of efficacy can be exhausting and lead one to evade the difficult tasks (Pajares, 1997).

Bandura (2006) has cautioned researchers attempting to measure self-efficacy in that there is no global measure of self-efficacy. An adequate analysis of the self-efficacy requires a detailed assessment of the level, generality, and strength of self-efficacy. The level of self-efficacy refers to variations across levels of task difficulty. The generality of the self-efficacy is concerned with the transfer of self-efficacy beliefs across activities. The strength of self-efficacy is measured by degrees of certainty that one can perform given tasks. These three dimensions of self-efficacy are measured using items that are task specific and vary in difficulty as well as degrees of confidence (Pajares, 1997; Zimmerman, 2000).

In the context of the teaching profession, teaching efficacy beliefs refer to teachers' judgements in their capabilities to conduct specific teaching tasks required to influence student learning (Dellinger, Bobbett, Olivier, & Ellett, 2008). It has been documented that teacher self-efficacy beliefs are powerful predictors of a variety of student- and teacher-related outcomes (Aldhafri, 2016; Garvis, 2013). As such, the measurement of teacher self-efficacy has received notable attention in educational research. Numerous instruments have been created to measure the construct. However, the majority of them were not consistent with Bandura's theoretical framework for measuring self-efficacy in terms of task specificity, thereby causing misinterpretation about the nature of self-efficacy itself (Tschannen-Moran, Hoy, & Hoy, 1998). Hence, Bandura (1997) developed Teacher Self-Efficacy Scale (TSS) consisting of 30 items divided into seven subscales: (a) efficacy to influence decision making, (b) efficacy to influence school resources, (c) instructional efficacy, (d) disciplinary efficacy, (e) efficacy to enlist parental involvement, (f) efficacy to enlist community involvement, and (g) efficacy to create a positive school environment. Not agreeing with Bandura's (1997) TSS's seven constructs conceptualizing the teachers' tasks, Tschannen-Moran and Woolfolk Hoy (2001) developed the Teacher Sense of Efficacy Scale (TSES) consisting of 24 items capturing three dimensions of teaching: (a) efficacy for student engagement, (b) efficacy for instructional strategies, and (c) efficacy for classroom management. Although satisfactory levels of reliability and validity were found for TSES compared to other teacher self-efficacy instruments, it represents a measure of the general teacher self-efficacy rather than a specific measure of self-efficacy related to the content being taught such as mathematics.

When considering mathematics teaching and learning, it has been documented, on one hand, that student mathematics self-efficacy is a predictor of mathematics student achievement which is in itself affected by teachers' practices (Pajares & Miller, 1994; Liu & Koirala, 2009). On the other hand, self-efficacy of mathematics teachers have

been found to affect their beliefs and practices related to mathematics instruction (Stipek, Givvin, Salmon, & MacGyvers, 2001). Thus, examining self-efficacy of mathematics teachers is important for enhancing student mathematics self-efficacy and academic achievement. To this end, researchers have been studying the measurement of teachers' self-efficacy with regard to mathematics.

Ernest (1989) described teachers' beliefs about mathematics as comprising of ideas about mathematics as a subject for study, the nature of mathematics teaching, and mathematics learning. As such, instruments developed to measure teacher self-efficacy with regard to mathematics capitulate two constructs: self-efficacy for teaching mathematics and mathematics self-efficacy. Self-efficacy for teaching mathematics refers to a teacher's belief regarding his or her ability to teach mathematics to others whereas mathematics self-efficacy refers to a teacher's belief regarding his or her ability to perform mathematical tasks (Kahle, 2008). An example of a measure of the mathematics self-efficacy includes The Mathematics Self-Efficacy Scale developed by Betz and Hackett (1983) by asking participants to rate their confidence in their ability to solve mathematical problems. The present study focuses on the self-efficacy for teaching mathematics.

A number of instruments have been cited in the literature as being used to assess self-efficacy of teaching mathematics. For example, Enochs, Smith, and Huinker (2000) developed the Mathematics Teaching Efficacy Beliefs Instrument (MTEBI) for pre-service elementary teachers. The instrument was developed as a result of the modification of the Science Teaching Efficacy Belief Instrument (STEBI). It consisted of 21 items measuring two constructs: (a) personal mathematics teaching efficacy (13 items) and (b) mathematics teaching outcome expectancy (8 items). Although the validity and reliability of the MTEBI have been supported by some studies (e.g., Alkhateeb, 2004; Cakiroglu, 2008; Cetinkaya & Erbas, 2011), Kieftenbeld, Natesan, and Eddy (2010) questioned its psychometric properties using item response theory and called for further validation of the instrument. Additionally, McGee (2012) questioned the construct validity of the MTEBI by arguing that the mathematics teaching outcome expectancy subscale is not theoretically a component of self-efficacy. As such, McGee (2012) developed The Self-Efficacy for Teaching Mathematics Instrument (SETMI) for in-service elementary mathematics teachers in the United States. It consisted of 37 items divided into three dimensions: (a) mathematics teaching self-efficacy (7 items), (b) self-efficacy for understanding elementary mathematics content (15 items), and (c) efficacy for teaching mathematics content (15 items). Although McGee (2012) reported that the SETMI is both valid and reliable, the author called for further validation of the instrument to meet the context specificity requirements of the measurement of self-efficacy. In response to this call, the present study aimed at developing and validating a scale for measuring self-efficacy of mathematics teachers by tailoring the content of the scale's items to match the mathematics curriculum of grades 5 to 10 in the Sultanate of Oman. The study utilized the structural models of the McGee's (2012) SETMI and the survey developed by Aerni (2008) to assess mathematics teachers' beliefs related to inclusive education. In addition, the study attempts to explore the construct validity the scores from the scale by testing group differences and relationships of the construct with

respect to gender, the taught grade level and desire to leave teaching profession, teaching experience, and teaching load.

Research Problem

The originality of this study is three-fold. Firstly, of increased interest to educators in Oman is the enhancement of student academic achievement in mathematics. This interest springs out of the results from international tests of academic achievement such as TIMSS which have shown that Omani student academic achievement in mathematics is much lower than student achievement in many Middle East countries (Al-Balushi, Ambusaidi, Al-Mherzi, Al-Harthi, Al-Saadi, &Al-Balushi, 2014). Therefore, researchers have become interested in factors that can enhance student achievement in mathematics. One of these factors that have been postulated as a potential predictor of mathematics student academic achievement is teachers' self-efficacy beliefs about mathematics (McGee, 2012). Thus, the research problem being considered in the current study is the development and datametric properties of a scale for measuring self-efficacy for teaching mathematics in grades 5 to 10 for teachers in Oman. Detailed information about such tool is helpful to those preparing and supervising mathematics teachers in the development of appropriate training programs supporting the enhancement of teaching practices of mathematics. Secondly, the statistics regarding the relationships between the components of the scale in terms of general versus discipline specific self-efficacy as well as the relationships with other variables are useful as part of the information educators use to decide whether to use general or discipline specific skills to foster the teaching of mathematics considering the contextual factors. Thirdly, this study adds to the body of knowledge regarding the measurement of self-efficacy for teaching mathematics.

Purposes of the study

The purposes of this study were to report on the development and datametric properties of a scale measuring the self-efficacy for teaching mathematics in grades 5 to 10 in Oman. The study would provide information regarding its construct validity by means of factor structure and group differences with respect to selected variables. The study would also present evidence concerning the reliability and quality of scale items. Additionally, the study would provide preliminary data regarding concurrent validity in terms of the correlations of the subscales' scores with each other.

METHOD

Sample

The participants in this study were 328 mathematics teachers randomly selected from five educational governorates in the Sultanate of Oman. They were teaching mathematics for grades 5 to 10 in public schools. The sample included 164 males and 164 females. The majority of them (96.4%) had an undergraduate degree in mathematics education and the remaining 3.6% had a postgraduate degree in education. Their teaching experience ranged from 2 to 26 years with an average of 12 years and a standard deviation of 4 years. Approximately three fourth of the participants had 1 to 5

workshops related to the mathematics teaching. The majority of the teachers (82%) were teaching in middle socioeconomic status schools, 11.6% in low socioeconomic status schools, and 6.4% in high socioeconomic status schools.

Procedures

After obtaining permission from the Ministry of Education and schools' principals, the data collection process took place in each school during a scheduled meeting. The teachers were informed that a study is being conducted to examine their confidence in teaching mathematics. They were informed that they were not obligated to participate in the study, and if they wished to participate, their responses would remain anonymous and confidential. Those who wished to participate were asked to respond to the scale described below.

Development of the Scale

The scale was developed within three stages. The first stage involved examining several teacher self-efficacy instruments that have been developed both generally and within contexts. Four instruments were used as a guide for the construction of items: The Teacher Self-Efficacy Scale (TSES) developed by Tschannen-Moran and Woolfolk Hoy (2001) to measure the general teacher efficacy expectations and outcome expectations; the Mathematics Teaching Efficacy Belief Instrument (MTEBI) developed by Enochs, Smith, and Huinker (2000) to measure general efficacy beliefs in teaching mathematics; Teaching Mathematics in Inclusive Settings (TMIS) developed by Aerni (2008) to measure efficacy beliefs in teaching mathematics in inclusive settings; and the Self-Efficacy for Teaching Mathematics Instrument (STEMI) developed by McGee (2012) to measure mathematics teaching self-efficacy for understanding mathematics content and efficacy for teaching mathematics content for elementary teachers. In this study, the items of the scale were constructed to balance between the generality and the specificity. Specifically, they were developed to measure teacher's efficacy in a specific context which was mathematics content of grades 5 to 10 in Oman.

The second stage involved analysing the main teaching and learning objectives of Omani mathematics curriculum content and textbooks in grades 5 to 10. The content was characterized according to the scientific knowledge, skills, and positive trends expected to be achieved. Moreover, the Omani educational outcomes were also studied such as: the development of communication skills, the application of new learning and teaching strategies, the adoption of modern technologies in the classroom situation, and the application of sophisticated strategies to assess and evaluate students' outcomes.

The third stage involved writing up the items. They were divided into three parts. The first part included 20 items measuring the general teaching self-efficacy addressing instructional strategies, assessment strategies, student engagement, and classroom management. The items were modified to be mathematics specific. The second part included 35 items measuring the teacher self-efficacy for understanding the mathematics content related to grades 5 to 10. The third part included 35 items measuring the teacher self-efficacy for teaching the mathematics content related to grades 5 to 10. The

participants were asked to rate their confidence in their ability to do the task described in each item on a 9-point Likert scale ranging from 1 (*not at all*) to 9 (*very high*).

The items were written in Arabic. The back-translation design (Hambleton, 2005) guided the adaption of the general efficacy items into Arabic. Accuracy of the translation was verified by the research team. Few editing modifications were made as a result of the translation.

Content validity of the items was established by presenting them to a panel of experts in the areas of mathematics education and educational measurement and psychology. They were asked to judge the clarity of wording and appropriateness of each item for the use with the targeted participants and its comprehensiveness and relevance to the construct being measured. Their feedback was used for refinement of the items. The consulted judges agreed that the items were clearly worded, appropriate for the participants, covering the mathematics content of grades 5 to 10, and relevant to the constructs being measured.

Data analysis

In the light of the aforementioned purposes of the study, the following statistical procedures were followed: (1) Quality of items were assessed by conducting item analysis in terms of item- endorsement, item discrimination, and alpha-if-item-deleted index; (2) The factor structure of the scale items was examined by principal components analysis; (3) The reliability was assessed by computing Cronbach's alpha internal consistency reliability estimates; (4) Pearson product-moment correlation coefficients were calculated between the subscales' scores; (5) Multivariate analysis of variance was conducted to examine differences on the self-efficacy for teaching mathematics with respect to gender, the taught grade level and the desire to leave teaching profession; and (6) Pearson product-moment correlation coefficients were calculated between the subscales' scores and years of teaching experience and teaching load defined as the number of weekly classes.

FINDINGS

Item analysis

Three types of item analysis were conducted to examine the quality of the items of the scale. These were item-endorsement index, item discrimination index, and alpha-if-item-deleted index. These indices show the contribution of each item to the measurement of the construct. The item-endorsement index indicates the percentage of individuals who endorsed the item. The item discrimination index indicates the relationship between the item's score and the total scale's score. The alpha-if-item-deleted index determines whether the score reliability of the scale would increase or decrease if the item was removed from the analysis. These indices were used together to determine whether the items are functioning satisfactorily in measuring the construct. Reasonable values for the item-endorsement index should be between .30 and .70 with an average of .50 whereas reasonable values for the item discrimination index should be greater than .40. Items with values outside these ranges should be considered for removal from the scale

because they do not maximize the information about the ability of the individuals (Cohen & Swerdlik, 2010).

In this study, the values were calculated for each subscale of the instrument: General teaching self-efficacy, self-efficacy for understanding the mathematics content, and self-efficacy for teaching the mathematics content. With respect to the items measuring general teaching self-efficacy, the values of the item-endorsement index ranged from .69 to .95 with an average of .84 and the values of the item discrimination ranged from .47 to .80 with an average of .72. With respect to the items measuring self-efficacy for understanding the mathematics content, the values of the item-endorsement index ranged from .86 to .95 with an average of .92 and the values of the item discrimination ranged from .78 to .92 with an average of .87. With respect to the items measuring self-efficacy for teaching the mathematics content, the values of the item-endorsement index ranged from .86 to .98 with an average of .93 and the values of the item discrimination ranged from .68 to .85 with an average of .76. The values of the alpha-if-item-deleted index did not show a negative contribution to the score reliability of each subscale. These results showed that the items are appropriate for the measurement of self-efficacy for teaching mathematics.

Factorial structure

The screening process of the data did not show concerns about normality, linearity, multicollinearity, and singularity. Inspection of the correlation matrix of the 90 items revealed that the correlations when taken overall were statistically significant as indicated by the Bartlett's test of sphericity, $\chi^2(4005) = 42556.06$, $p < .001$. Kaiser's measure of sampling adequacy (MSA) fell within an acceptable range (values of .60 and above) with a value of .94. Each item also exceeded the threshold value (.60) of MSA. Finally, most of the partial correlations were small as indicated by the anti-image correlation matrix. These measures all led to the conclusion that the set of 90 items of the self-efficacy for teaching mathematics was appropriate for principal components analysis (PCA).

Teachers' responses to the 90 items of the self-efficacy for teaching mathematics were submitted to PCA to identify their underlying dimensions. No particular number of dimensions was hypothesized and the criterion was set to eigenvalues greater than one (Tabachnick & Fidell, 2001). The initial unrotated PCA resulted in a factor model of four dimensions as suggested by the screen plot and eigenvalues exceeding unity. However, based on its pattern of factor loadings, this unrotated factor model was theoretically less meaningful and difficult to interpret. Therefore, the analysis proceeded to rotate the factor matrix orthogonally with Varimax rotation to achieve a simple and theoretically more meaningful solution.

This analysis yielded three factors as suggested by the eigenvalue rule and scree plot. Together the three factors accounted for 64.65% of the total variance. All items loaded ≥ .43 on their primary factor. The first factor accounted for 40.77% of the variance (eigenvalue = 36.69) and consisted of 35 items with factor loadings ranging between .69 and .93. According to the content of its items and in light of the literature (Aerni, 2008;

Enochs, Smith, & Huinker, 2000; McGee, 2012; Tschannen-Moran & Woolfolk Hoy, 2001), this factor was labelled as *self-efficacy for understanding the mathematics content* because its items focused on teachers' confidence in their ability to understand mathematics content specific to grades 5 to 10. The second factor accounted for 17.37% of the variance (eigenvalue = 15.64) and consisted of 35 items with factor loadings ranging between .64 and .79. According to the content of its items and in light of the literature (Aerni, 2008; Enochs, Smith, & Huinker, 2000; McGee, 2012; Tschannen-Moran & Woolfolk Hoy, 2001), this factor was named as *self-efficacy for teaching the mathematics content* because its items focused on teachers' confidence in their ability to teach students mathematics content specific to grades 5 to 10. The third factor accounted for 6.51% of the variance (eigenvalue = 5.86) and consisted of 20 items with factor loadings ranging between .43 and .76. According to the content of its items and in light of the literature (Aerni, 2008; Enochs, Smith, & Huinker, 2000; McGee, 2012; Tschannen-Moran & Woolfolk Hoy, 2001), this factor was labelled as *general teaching self-efficacy* because its items focused on general rather than specific aspects of teachers' efficacy concerning the use of instructional strategies, assessment strategies, student engagement, and classroom management when teaching mathematics.

Reliability

Internal consistency coefficients for subscales' scores of self-efficacy for understanding the mathematics content, self-efficacy for teaching the mathematics content, and general teaching self-efficacy were .99, .98, and .95 as indicated by Cronbach's alpha, respectively. Internal consistency coefficient for the total scale's scores was .98. These values indicate that the Self-Efficacy for Teaching Mathematics Scale have a strong score reliability.

Concurrent validity

Concurrent validity of the scale was examined by correlating self-efficacy for understanding mathematics content and self-efficacy for teaching the mathematics content to the general teaching self-efficacy. Additionally, subscales' scores were correlated with the total scale scores. Measures of self-efficacy for understanding the mathematics content, self-efficacy for teaching the mathematics content, and general teaching self-efficacy were constructed by averaging the items on each factor. Table 1 presents intercorrelations, means, and standard deviations for the measures. On average, the participants reported high levels of self-efficacy for understanding mathematics content, self-efficacy for teaching the mathematics content, and general teaching self-efficacy. As shown in Table 1, the general teaching self-efficacy had statistically significant positive correlations with the self-efficacy for understanding mathematics content and self-efficacy for teaching the mathematics content. The relationship of the general teaching self-efficacy to the self-efficacy for teaching mathematics content was stronger than to the self-efficacy for understanding mathematics content. All subscales correlated positively and strongly with the total scale.

Table 1
Intercorrelations, Means, and Standard Deviations for Self-Efficacy for Understanding Mathematics Content, Self-Efficacy for Teaching the Mathematics Content, General Teaching Self-Efficacy, and total Scale Scores (N = 328)

Variable	1	2	3	4	M	SD
1. SEUMC	-	..43***	.30***	.83***	8.24	1.18
2. SETMC		-	.58**	.82***	8.32	.89
3. GTSE			-	.69***	7.57	1.02
4. Total scale				-	8.12	.82

Note. SEUMC = Self-efficacy for understanding mathematics content; SETMC = Self-efficacy for teaching the mathematics content; GTSE = General teaching self-efficacy.
***$p < .001$.

Group differences

A multivariate analysis of variance (MANOVA) was conducted to examine differences on the self-efficacy for teaching mathematics with respect to gender, the taught grade level and the desire to leave teaching profession. Table 2 presents means and standard deviations for self-efficacy for understanding mathematics content, self-efficacy for teaching the mathematics content, and general teaching self-efficacy based on gender, the taught grade level and the desire to leave teaching profession. Table 3 summarizes results of the MANOVA. As shown in Table 3, there was a statistically significant multivariate interaction effect on the self-efficacy for teaching mathematics between gender and the desire to leave the teaching profession. This interaction effect accounted for 2.7% of the variability in the self-efficacy for teaching mathematics.

Table 2
Means and Standard Deviations for Self-Efficacy for Understanding Mathematics Content, Self-Efficacy for Teaching the Mathematics Content, and General Teaching Self-Efficacy based on Gender, the Taught Grade Level and the Desire to Leave Teaching Profession

Variable	n	SEUMC		SETMC		GSET	
		M	SD	M	SD	M	SD
Gender							
Male	164	8.05	1.25	8.08	1.07	7.18	1.10
Female	164	8.42	1.08	8.57	.56	7.95	.77
Grade level							
Five	44	8.40	1.00	8.29	.78	7.85	.76
Six	43	8.26	1.05	8.23	.97	7.63	1.01
Seven	60	8.42	.95	8.38	.85	7.45	1.15
Eight	59	7.99	1.36	8.24	.99	7.60	1.01
Nine	61	8.15	1.27	8.31	.85	7.44	1.10
Ten	61	8.24	1.31	8.46	.88	7.52	.98
Leave teaching							
No	155	8.26	1.14	8.36	.72	7.62	.89
Yes	173	8.22	1.22	8.29	1.01	7.52	1.13

Note. SEUMC = Self-efficacy for understanding mathematics content; SETMC = Self-efficacy for teaching the mathematics content; GTSE = General teaching self-efficacy.

Table 3
Summary of MANOVA for the Effect of Gender, the Taught Grade Level and the
Desire to Leave Teaching Profession on the Self-Efficacy for Teaching Mathematics

Source	*Wilks' lamda*	*F*	*Hypothesis df*	*Error df*	*p-value*	*Partial* η^2
Gender	.86	16.48	3	302	.000	.14
Grade	.93	1.43	15	834	.128	-
Leave teaching	.99	.60	3	302	.615	-
Gender×Grade	.96	.87	15	834	.602	-
Gender×Leave	.97	2.83	3	302	.039	.027
Grade×Leave	.97	.71	15	834	.778	-
Gender×Grade×Leave	.93	1.45	15	834	.118	-

The univariate analysis showed statistically significant gender differences favouring females on the self-efficacy for understanding mathematics content; $F(1, 304) = 5.94$, $p = .015$, partial $\eta^2 = .019$; and general teaching self-efficacy $F(1, 304) = 46.59$, $p = .000$, partial $\eta^2 = .133$. Also, there was a statistically significant interaction effect on the self-efficacy for teaching the mathematics content between gender and the desire to leave the teaching profession; $F(1, 304) = 7.18$, $p = .008$, partial $\eta^2 = .023$. Figure 1 displays the interaction effect on the self-efficacy for teaching the mathematics content between gender and the desire to leave the teaching profession. As shown in Figure 1, the difference between males and females in the level of self-efficacy for teaching the mathematics content for those who want to leave the teaching profession was higher than that for those who do not want to leave the teaching profession.

Figure 1. Interaction effect between gender and desire to leave teaching profession on the self-efficacy for teaching the mathematics content

Further validation: Associations with teaching experience and teaching load

An additional validation of the measures was conducted by exploring the relationships of teaching experience and weekly teaching load to the subscales of the self-efficacy for teaching mathematics. Pearson product-moment correlation analyses showed no statistically significant relationships exist between the subscales of the self-efficacy for teaching mathematics and teaching experience and weekly teaching load.

DISCUSSION

Given the unsatisfied performance of Omani students in mathematics, the teaching and learning of mathematics continue to be of national concern (Al-Balushi et al., 2014). As such, identification of factors that might contribute to the enhancement of mathematics education becomes certainly warranted. One of these factors that have received increased credence in the literature is teacher self-efficacy. As stated by Tschannen-Moran and Wolfolk Hoy (2001), "Teacher efficacy has proved to be powerfully related to many meaningful educational outcomes such as teachers' persistence, enthusiasm, commitment and instructional behavior, as well as student outcomes such as achievement, motivation, and self-efficacy beliefs" (p. 783). With increased interest in research into teacher self-efficacy beliefs, the measurement of this construct and its culturally cross-validation have attracted more attention from educational researchers. Building on self-efficacy measures developed by Tschannen-Moran and Woolfolk Hoy (2001), Enochs et al. (2000), Aerni (2008), and McGee (2012), the Self-Efficacy for Teaching Mathematics Scale was developed in this study to measure teachers' judgements in their ability to teach mathematics content related to grades 5 to 10 in Oman. Principal components analyses yielded a scale with 90 items divided into three subscales: self-efficacy for understanding mathematics content (35 items), self-efficacy for teaching the mathematics content (35 items), and general teaching self-efficacy (20 items). This structure agrees with the structure found in McGee's (2012) study.

The three derived subscales demonstrated good score reliability and validity. Evidence for reliability was based on Cronbach's alpha coefficients, which were all greater than .90. Evidence for validity was based on factorial structure in which all items loaded greater than .40 on their primary factors. An additional construct validity evidence was based on empirical testing of group differences which showed the ability of the scale to differentiate between teachers in terms of their self-efficacy for teaching mathematics with medium to large effect sizes based on gender and the desire to leave the teaching profession . Gender differences on the mathematics teaching self-efficacy beliefs are consistent with the findings reported by Cetinkaya and Erbas (2011) who found that female teachers held higher levels of personal mathematics teaching efficacy than males. Also, the findings support those reported by Klassen and Chiu (2010) who showed a significant relationship between teachers' self-efficacy and job satisfaction. Yet, the present study adds to the previous studies in the interaction effect of gender and the desire to leave the teaching profession on the mathematics teaching efficacy beliefs. Additionally, the current study findings agree with those reported by Cetinkaya and Erbas (2011) with regard to the non-significant effect of teaching experience on the

mathematics teaching efficacy beliefs. Nonetheless, more research is needed to strengthen the construct validity of the scale by means of group differences.

Furthermore, results of the concurrent validity analysis revealed moderate correlations between the subscales thereby suggesting that the subscales assess related, yet unique aspects of the self-efficacy for teaching mathematics. These results lend support to using the subscales separately. In addition, results of the analysis at the item-level further strengthened the internal structure of the scale. Also, results of the item analysis suggested administering the items in order of average confidence level. Taken together, these results support the use of the Self-Efficacy for Teaching Mathematics Scale as an assessment tool for educational and research purposes. The information obtained from the three subscales can lead to an understanding of the relative strengths and weaknesses of the teachers' skills in teaching mathematics. This kind of information can help in developing professional training programs for teachers targeting the enhancement of specific skills and domains related to the teaching of mathematics.

Continued research is needed to increase the evidence of the scale reliability and validity. First, the sample size did not permit an examination of the structural characteristics of the scale's items using confirmatory factor analysis. It is possible that the factors extracted in the present study about self-efficacy for understanding mathematics content and self-efficacy for teaching the mathematics content could be related to a higher-order factor. Future research might consider evaluating this possibility statistically. Second, the sample size might have also lacked the power to detect statistically significant group differences on the self-efficacy for teaching mathematics with respect to teaching experience and teaching load. Future research might examine such differences. Third, the scale needs to be validated for teachers teaching mathematics in grades other than grades 5 to 10 considered in the current study. Fourth, the selectivity of the teachers in this study limits the generalizability of the findings. Specifically, the selected teachers in this study were exposed to many sources of efficacy information to reinforce high levels of efficacy. Most of them had 1 to 5 training workshops related to the mathematics teaching and the majority of them were teaching in middle socioeconomic status schools. As such, the scale needs to be validated in larger as well as diverse socioeconomic school settings. Finally, additional studies are needed to assess the predictive validity of the scale in terms of teacher effectiveness in teaching mathematics and subsequent student performance.

REFERENCES

Aerni, P. W. (2008). *Teacher self-efficacy and beliefs for teaching mathematics in inclusive settings* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3353198).

Al-Balushi, S., Ambusaidi, A. K., Al-Mherzi, R., Al-Harthi, I. S., Al-Saadi, K. A., & Al-Balushi, M. (2014). *Students' performance in TIMSS test in Oman: Effect of cognitive and metacognitive variables and effectiveness of an inquiry-based mobile e-formative assessment package*. A proposal of a research project funded by The Research Council   (TRC)-Oman.   Retrieved   from

https://www.researchgate.net/publication/301496050_Students'_performance_in_TIMSS_test_in _Oman_Effect_of_cognitive_and_metacognitive_variables_and_effectiveness_of_an_inquiry- based_mobile_e-formative_assessment_package_tathyr_almtghyrat_almrfyt_wfwq_a.

Aldhafri, S. (2016). Predicting pre-school teachers' burnout levels through their efficacy beliefs in the Sultanate of Oman. *Journal of Arabic Childhood, 66*, 105-115.

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.

Bandura, A. (2006). Guide for constructing self-efficacy scales. In F. Pajares & T. Urdan (Eds.) *Adolescence and education: Vol. 5. Self-efficacy and adolescence* (pp. 307 – 337). Greenwich, CT: Information Age.

Bandura, A. (1997). *Self-Efficacy: The Exercise of Control*. New York: W. H. Freeman and Company.

Betz, N. E., & Hackett, G. (1983). The relationship of mathematics self-efficacy expectations to the selection of science-based college majors. *Journal of Vocational Behavior, 23*, 329 – 345. doi: 10.1016/0001-8791(83)90046-5.

Cakiroglu, E. (2008). The teaching efficacy beliefs of pre-service teachers in the USA and Turkey. *Journal of Education for Teaching, 34*, 33 – 44. doi: 10.1080/02607470701773457.

Cetinkaya, B., & Erbas, A. K. (2011). Psychometric properties of the Turkish adaptation of the mathematics teacher efficacy belief instrument for in-service teachers. *The Spanish Journal of Psychology, 14*, 956 – 966. doi: 10.5209/rev_SJOP.2011.v14.n2.41.

Cohen, R. J., & Swerdlik, M. E. (2010). *Psychological Testing and Assessment: An Introduction to Tests and Measurement* (7th ed.). Boston: McGraw-Hill.

Dellinger, A. B., Bobbett, J. J., Olivier, D. F., & Ellett, C. D. (2008). Measuring teachers' self-efficacy beliefs: Development and use of the TEBS-Self. *Teaching and Teacher Education, 24*, 751 – 766. doi: 10.1016/j.tate.2007.02.010.

Enochs, L. G., Smith, P. L., & Huinker, D. (2000). Establishing factorial validity of the Mathematics Teaching Efficacy Beliefs Instrument. *School Science and Mathematics, 100*, 194 – 202. doi: 10.1111/j.1949-8594.2000.tb17256.x.

Ernest, P. (1989). *Mathematics Teaching: The State of the Art*. London: The Falmer Press.

Garvis, S. (2013). Beginning generalist teacher self-efficacy for music compared with maths and English. *British Journal of Music Education, 30*, 85 – 101. doi: 10.1017/S0265051712000411.

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.

Khale, D. K. B. (2008). *How elementary school teachers; mathematical self-efficacy and mathematics teaching self-efficacy relate to conceptually and procedurally oriented teaching practices* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3313038).

Klassen, R., & Chiu, M. M. (2010). Effects on teachers' self-efficacy and job satisfaction: Teacher gender, years of experience, and job stress. *Journal of Educational Psychology, 102*, 741 – 756. doi: 10.1037/a0019237.

Liu, X., & Koirala, H. (2009). The effect of mathematics self-efficacy on mathematics achievement of high school students. *NERA Conference Proceedings paper 30*. Retrieved from http://digitalcommons.ucom.edu/nera_2009/30.

McGee, J. R. (2012). *Developing and validating a new instrument to measure the self-efficacy of elementary mathematics teachers* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3510218).

Pajares, M. F. (1997). Current directions in self-efficacy research. In M. Maehr & P. R. Pintrich (Eds.) *Advances in motivation and achievement, 10* (pp. 1 – 49). Greenwich, CT: JAI Press. Retrieved from http://www.uky.edu/~eushe2/Pajares/effchapter.html.

Pajares, F., & Miller, M. D. (1994). Role of self-efficacy and self-concept beliefs in mathematical problem solving: A path analysis. *Journal of Educational Psychology, 86*, 193 – 203. doi: 10.1037/0022-0663.86.2.193.

Stipek, D. J., Givvin, K. B., Salmon, J. M., & MacGyvers, V. L. (2001). Teachers' beliefs and practices related to mathematics instruction. *Teaching and Teacher Education, 17*, 213 – 226. doi: 10.1016/S0742-051X(00)00052-4.

Tabachnick, B. G., & Fidell, L. S. (2001). *Using Multivariate Statistics* (4th ed.). Needham Heights, MA: Allyn & Bacon.

Tschannen-Moran, M., & Woolfolk Hoy, A. W. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and Teacher Education, 17*, 783 – 805. doi: 10.1016/S0742-051X(01)0036-1.

Tschannen-Moran, M., Hoy, A. W., & Hoy, W. K. (1998). Teacher efficacy: Its meaning and measure. *Review of Educational Research, 68*, 202 – 248. doi: 10.3102/00346543068002202.

Zimmerman, B. J. (2000). Self-efficacy: An essential motive to learn. *Contemporary Educational Psychology, 25*, 82 – 91. doi: 10.1006/ceps.1999.1016.

Turkish Abstract

Umman Sultanlığında Öğretmenlerin Matematik Öğretimine Yönelik Öz-Yeterlik Ölçüm Ölçeğinin Geliştirilmesi ve Onaylanması

Bu çalışmada Umman'daki 5'ten 10. sınıfa kadar olan öğretmenlerin matematik öğretimindeki öz yeterliliğini ölçmek için bir ölçek geliştirilmiştir. Katılımcılar, Umman Sultanlığı'nda beş eğitim bölgesi arasından rastgele seçilen 328 matematik öğretmeni olarak belirlenmiştir. Ölçeğin faktörel yapısı, üç alt ölçek ortaya koymuştur: matematik içeriğini anlama için öz yeterlik, matematik içeriğini öğretmek için öz yeterlik ve genel öğretim öz yeterliği. Üç alt ölçek, yeterli bir şekilde güçlü iç tutarlılık ve yapı geçiciliği ve eşzamanlı geçerlik kanıtları göstermiştir. Ölçek hem eğitim hem de araştırma amaçları için potansiyel kullanımlara sahiptir.

Anahtar Kelimeler: öz-yeterlilik, öğretimin etkililiği, matematiğin etkililiği, güvenilirlik, ölçek geliştirme

French Abstract

Développement et Validation d'une Échelle pour Mesurer Mathématiques Enseignant Auto-efficacité pour Professeurs dans le Sultanat de l'Oman

Une échelle pour mesurer l'auto-efficacité pour enseigner des mathématiques dans des catégories 5 à 10 a été développée dans cette étude pour des professeurs à l'Oman. Les participants étaient 328 professeurs de mathématiques aléatoirement choisis de cinq gouvernorats éducatifs dans le Sultanat de l'Oman. La structure de factorielle de l'échelle a révélé trois sous-échelles : l'auto-efficacité pour comprendre le contenu de mathématiques, l'auto-efficacité pour enseigner le contenu de mathématiques et le général enseignant l'auto-efficacité. La trois sous-échelles a montré la coherence interne forte et l'évidence suffisante de validité de construction et la validité simultanée. L'échelle a des utilisations potentielles tant pour éducatif que des buts de recherche.

Mots Clés: l'auto-efficacité, enseignant l'efficacité, l'efficacité de mathématiques, la validité, la fiabilité, pèse le développement

Arabic Abstract

تطوير وتثبيت مقياس لقياس تعليم الرياضيات فعالية ذاتية للمعلمين في سلطنة عمان

تم تطوير مقياس لقياس الكفاءة الذاتية لتدريس الرياضيات في الصفوف من 5 إلى 10 في هذه الدراسة للمعلمين في عمان. وقد تم اختيار 328 معلما رياضياتيا من خمس محافظات تعليمية في سلطنة عمان. وكشف الهيكل المعياري للمقياس عن ثلاثة مستويات فرعية: الكفاءة الذاتية لفهم محتوى الرياضيات، والكفاءة الذاتية لتدريس محتوى الرياضيات، والكفاءة الذاتية التدريس العام. وأظهرت الجداول الفرعية الثلاثة اتساقا داخليا قويا وأدلة كافية على صحة البناء والصلاحية المتزامنة. وللمقياس استخدامات محتملة لكل من الأغراض التعليمية والبحثية.

الكلمات الرئيسية: الكفاءة الذاتية، فعالية التدريس، فعالية الرياضيات، والصلاحية، والموثوقية، وتطوير نطاق

German Abstract
Entwicklung und Validierung einer Maßstab für die Messung der Mathematik Lehre Selbstwirksamkeit für Lehrer im Sultanat von Oman

Eine Skala zur Messung der Selbstwirksamkeit für den Mathematikunterricht in den Klassen 5 bis 10 wurde in dieser Studie für Lehrer in Oman entwickelt. Die Teilnehmer waren 328 Mathematiklehrer, die zufällig aus fünf Bildungsgouvernements im Sultanat von Oman ausgewählt wurden. Die faktorische Struktur der Skala zeigte drei Subskalen: Selbstwirksamkeit für das Verständnis des Mathematikinhalts, die Selbstwirksamkeit für den Unterricht des Mathematikinhalts und die allgemeine Lehre Selbstwirksamkeit. Die drei Subskalen zeigten eine starke innere Konsistenz und ausreichende Beweise für die Konstruktvalidität und die gleichzeitige Gültigkeit. Die Skala hat Potenziale für pädagogische und Forschungszwecke.

Schlüsselwörter: selbstwirksamkeit, lehre wirksamkeit, mathematik wirksamkeit, gültigkeit, zuverlässigkeit, skalenentwicklung

Malaysian Abstract
Pembangunan dan Pengesahan Skala Pengukuran Pengajaran Matematik Efikasi Kendiri Guru di Kesultanan Oman

Skala pengukuran efikasi kendiri untuk mengajar matematik dalam gred 5 hingga 10 telah dibangunkan dalam kajian ini untuk guru-guru di Oman. Para peserta terdiri daripada 328 orang guru matematik yang dipilih secara rawak daripada lima governorates pendidikan di Kesultanan Oman. Skala struktur faktorial mendedahkan tiga subskala: keberkesanan diri untuk memahami kandungan matematik, keberkesanan diri untuk mengajar kandungan matematik, dan pengajaran biasa efikasi-kendiri. Tiga subskala menunjukkan ketekalan dalaman yang kukuh dan keterangan yang mencukupi mengenai kesahihan konstruk dan kesahan serentak. skala ini mempunyai kegunaan yang berpotensi untuk kedua-dua tujuan pendidikan dan penyelidikan.

Kata Kunci: efikasi kendiri, efikasi pengajaran, efikasi matematik, kesahan, kebolehpercayaan, pembangunan skala

Russian Abstract
Разработка и Аттестация Шкалы Измерения Математики Преподавания Самоподготовки для Учителей в Султанате Оман

В этом исследовании для учителей в Омане была разработана шкала оценки самоэффективности для преподавания математики в 5-10 классах. Участниками были 328 учителя математики, случайным образом отобранных из пяти образовательных мухафаз в Султанате Оман. Факториальная структура шкалы выявила три подшкалы: самоэффективность для понимания содержания математики, самоэффективность для преподавания математического содержания и самоэффективность общего преподавания. Три подшкалы показали сильную внутреннюю согласованность и достаточные доказательства достоверности конструкции и параллельной действительности. Шкала имеет потенциальное применение как для образовательных, так и для исследовательских целей.

Ключевые Слова: самоэффективность, эффективность преподавания, эффективность математики, обоснованность, надежность, развитие шкалы