MODESTUM EURASIA Journal of Mathematics, Science and Technology Education, 2023, 19(9), em2321
ISSN:1305-8223 (online)
OPEN ACCESS Research Paper https://doi.org/10.29333/ejmste/13481

# Adaptation of the sources of the mathematics self-efficacy scale for Oman: A validation study

Khalid Saif Salim Al Umairi [1*], Umi Kalsum Mohd Salleh [1], Hutkemri Zulnaidi [2]

[1] Department of Curriculum and Instructional Technology, Faculty of Education, Universiti Malaya, Kuala Lumpur, MALAYSIA
[2] Department of Mathematics and Science Education, Faculty of Education, Universiti Malaya, Kuala Lumpur, MALAYSIA

Received 03 April 2023 ▪ Accepted 20 May 2023

**Abstract**

The current study aims to adapt sources of middle school mathematics scale developed by Usher and Pajares (2009) to the Omani context. The study sample consisted of 700 students (379 girls and 321 boys) from the eighth grade in middle schools in Oman. The results of this study showed that sources of middle school mathematics self-efficacy scale (SMES) (adaption) resemble the original version of which developed by Usher and Pajares (2009). Confirmatory factor analysis showed good fit indices, with $\chi^2$=767.8, df=224, comparative fit index=.928, standardized root mean square residual=.0478, and root mean square error approximation=.059. The scale also displayed a robust internal consistency–Cronbach's alpha coefficient was above 0.8 and for its subscales as well. The indices of convergent and divergent validity for this scale were all significant ($p<0.001$). Therefore, SMES is psychometrically sound and can be applied to the Omani context.

**Keywords:** sources of self-efficacy, mathematics, scale, middle school

## INTRODUCTION

Mathematics achievement is a critical issue around the world. Scholars have tried to determine the best way to improve students' achievement in mathematics. To this end, several kinds of studies have been conducted in which self-efficacy was found to be one of the essential factors of a student's achievement in mathematics (Liu & Koirala, 2009; Recber et al., 2018; Skaalvik et al., 2015).

Additionally, the literature has revealed that a robust, positive relationship exists between mathematics self-efficacy and mathematics achievement (Ayotola & Adedeji, 2009; Evans, 2015; Meral et al., 2012). Students with high degrees of self-efficacy were found to be more precise in mathematic task and showed greater diligence in difficult tasks when compared to students with low self-efficacy (Ayotola & Adedeji, 2009). Self-efficacy influences the effort taken to complete a task and the persistence in focusing on that task, as well as academic motivation.

Thus, self-efficacy plays a significant role in students' academic achievement. However, according to the literature, most studies conducted on self-efficacy and on its role in achievement have been conducted in western contexts (Gao, 2019). Additionally, there is no scale of sources on mathematics self-efficacy in the context of Oman or Arab context, where "the strength and influence of the sources differ as a function of contextual factors such as gender, ethnicity, academic ability, and academic domain" (Usher & Pajares, 2008, p. 751).

Therefore, the current study will bridge the gap in the body of knowledge by translating and validating the sources of middle school mathematics self-efficacy scale (SMES), which was constructed by Usher and Pajares (2009) in the Omani context. Furthermore, this study was as one of the recommendations of Usher and Pajares (2009) to examine the validity of the items of SMES across different contexts and domains.

Over the last 34 years, scholars have applied the concept of self-efficacy in various fields of education to clarify a broad variety of human performances, such as academic achievement (Artino, 2012). Self-efficacy is considered as one of the major factors impacting a learner's cognition; it influences their actions and behaviors and guides their education. Moreover, the feeling of self-efficacy, that is a person's faith in their ability to complete a particular mission, encourages a


✉ khalid.alumairi@gmail.com **(*Correspondence)** ✉ umi_salleh@um.edu.my ✉ hutkemri@um.edu.my

**Contribution to the literature**

- Providing evidence for reliability and validity of SMES as a tool to measure sources of self-efficacy of students in mathematics in Oman context in light of the theory of Bandura (1977) is important.
- SMES has been used in different contexts for instance, Indonesia, Singapore, Taiwan, the USA, and Turkey. As well as it has used in different domains. SMES is a popular subject in the field.
- This study will assist scholars to understand the self-efficacy in Arabic context and it allow different scholars to conduct more studies on self-efficacy and its sources.

person to continue on a mission, become more actively engaged in it, and work in a harder and lengthier manner when pursuing that goal (Hendricks, 2016).

## THEORETICAL FRAMEWORK

Self-efficacy has garnered adequate interest in the area of learning research. Studies have demonstrated that it can be used to predict a learner's academic achievements in academic fields and levels (Usher & Pajares, 2008). Moreover, self-efficacy has been demonstrated to forecast a learner's college major and job selections (Brown & Lent, 2006, as cited in Usher & Pajares, 2008) and is related to main motivation structures, such as causal attributions, self-concept, optimism, achievement goal orientation, academic help-seeking, anxiety, and value (Usher & Pajares, 2008). It refers to specific activities and tasks in which people feel competent as opposed to a more general judgement about their abilities. (Van Dinther et al., 2011). Self-efficacy builds our emotions. To understand how people feel confident about their actions, it is necessary to investigate how self-efficacy functions. With regard to belief, a feeling of competence facilitates cognitive processes and performances in a range of settings that involve decision-making, goal-setting, and academic achievement (Husain, 2014).

The level of self-efficacy affects the way we think and how we feel. When self-efficacy beliefs are extremely high, people enjoy activities peacefully (Kumar & Lal, 2006). Self-efficacy beliefs also affect how well individuals motivate themselves and persist in facing obstacles that may arise from the targets they set for themselves, the resulting expectations, and the causative source for their success and failure. Self-regulation of emotional states is greatly influenced by beliefs about how people can cope. This influences the quality of their emotional well-being and their susceptibility to anxiety and despair (Bandura, 2012).

Furthermore, self-efficacy is a part of the social-cognitive theory of personality (Bandura, 1994). The latter argues that education takes place in social conditions with a dynamic and mutual interaction between the individual, their surroundings, and their actions (Zhou & Brown, 2015). According to the beliefs of social theorists, education is affected by a person's self-efficacy (Kehoe, 2013). In particular, an individual's beliefs in their efficiency affects them either negatively or positively in self-debilitating ways or self-enabling ways(Van Dinther et al., 2011).

Self-efficacy beliefs have four sources from which students also derive their beliefs. These sources are mastery experiences, vicarious experiences, verbal persuasion and physiological reactions or states. Self-efficacy is what learners deduce based on the knowledge they acquire from these sources. It is the judgement of their capability to be successful in a particular mission or a set of associated missions. By understanding and methodically utilizing these sources, educators can leave an impact on a learner's self-efficacy (Margolis & McCabe, 2006). According to Bandura (2012), the theory of self-efficacy hypothesizes that people obtain the knowledge to assess their efficacy beliefs from four main sources:

(a) mastery experience,

(b) vicarious experience,

(c) verbal persuasion, and

(d) physiological arousal.

### Mastery Experiences

Mastery experiences are considered to be a powerful source of self-efficacy (Britner & Pajares, 2006; Husain, 2014; Tschannen-Moran & Hoy, 2007). Bandura (1994) pointed out that the extremely efficient method of establishing a robust feeling of efficacy is via mastery experiences. Wherein, mastery experience is defined as individual's explanations of their personal earlier, real experiences accomplishing a specific mission. Mastery experiences demonstrate power, especially when people overcome difficulties or succeed in difficult tasks (Bandura, 1997, as cited in Usher & Pajares, 2008).

### Vicarious Experiences

Vicarious experiences of social models are another method by which beliefs of self-efficacy are established and boosted (Bandura, 1994). They are defined as individuals' assessments of their abilities with respect to the accomplishments of others (Chang, 2006). People learn to succeed by observing the success of a role model. Observing the success of others prompts people to succeed, whereas observing their failure leads to a decrease in self-efficacy beliefs (Bandura, 1994). Vicarious or indirect experiences are deemed as the second source of self-efficacy in which the factor that

helps create efficacy within a person is the vision that individuals on the same level as themselves succeed through sustained effort; this leads to an increase in the person's beliefs that they can master the same activities and be successful. In contrast, watching others' failures despite diligent attempts reduces the observer's judgement of their own efficacy and reduces their attempts (Bandura, 1994). Thus, the effect of modelling on recognized self-efficacy is based on its perceived similarity to the models, where whenever the assumed resemblance increases, the model's successes and failures are more persuasive (Bandura, 1994).

### Social Persuasion

The third vehicle for boosting individuals' self-efficacy beliefs is social persuasion, which convinces them that they have what it takes to succeed. Verbal persuasion is more widely used in education systems, as it is often used to convince students to believe in their abilities to cope with difficult situations (Artino, 2012). Social persuasion, which involves being vulnerable to the verbal and non-verbal judgements given by people, is also a crucial source of knowledge. Effective persuaders should work to develop the learner's beliefs in their abilities, while also guaranteeing that the predicted success is achievable (Britner & Pajares, 2006). Verbal persuasion refers to "the verbal judgments that others provide" (Chang, 2006, p. 55). Performing a definite task because of encouragement or persuasion from others is called social persuasion. For example, when someone convinces another person to perform some tasks, the person who is convinced by another tends to believe that they can perform this task.

### Physiological Arousal

The fourth source of self-efficacy involves physiological and emotional states. Individuals with a great feeling of efficacy see their status of sentimental stimulation as an invigorating performance easer, whereas those who are stressed by their self-belief consider their stimulation to be weakening (Bandura, 1994). People are also partly dependent on their physical and emotional states to judge their abilities. They explain their anxiety responses and nervousness as a sign of weakness that leads to poor performance. In actions such as potency and endurance, individuals judge their tiredness, soreness, and worries as signs of physical frailty. Attitude also influences individuals' judgements of their efficacy. While a positive attitude improves self-efficacy beliefs, a degrading attitude weakens it. Fourth method of changing self-belief of efficacy is to lessen an individual's anxiety responses and change their bad emotional tendencies and misinterpretation of their physical states (Bandura, 1994).

Physiological factors, such as fear, anxiety, stimulation and attitude, provide information regarding one's efficacy beliefs. Learners assess their level of trust based on the emotional state they are in as they observe or participate in an act. They are further likely to anticipate success when they encounter pleasant excitement than when they experience high levels of nervousness, pressure or strain concerning a particular action or sphere. Bad physical states–or those understood as bad–could hinder performance and increase the likelihood of a low outcome, thereby leading to poorer levels of self-efficacy. As a result, individuals also differ in degree to which they concentrate on their inner status and degree to which they are predisposed to relate good or bad results to those statuses. The current level of self-efficacy, the complexity of the mission at hand, and previous experiments in comparable positions also influence the affective states, the explanation of physiological states, and the impact they have on self-efficacy. Physiological states do not differ from earlier sources, where it is the explanation of physiological states that gives to self-efficacy (Britner & Pajares, 2006).

### SMES

Albert Bandura introduced the construct of self-efficacy beliefs in the last century, specifically the 1970s (Van Dinther et al., 2011). Since Bandura (1977) introduced the construct of self-efficacy and the problems of accurate measurement of self-efficacy based on Bandura's (1977) theory is still there between scholars. Accordingly, the scholars conducted the first debate on the issue in 1996 at one of the annual meetings of the American Educational Research Association held in New York, that is *Measuring and Mismeasuring Self-Efficacy: Dimensions, Problems, and Misconceptions Conference* (Echeverría Castro et al., 2020). According to that, Usher and Pajares (2009) made up their minds to make a critical review of the state of knowledge of the sources of school self-efficacy. They found that during the critical review in the measurement part, specifically the measurement of self-efficacy, there were a diversity of instruments that measure the four sources of self-efficacy or only some of them. They detected forms that do not resemble what is founded in the theory, as well as test items that contradict Bandura's (1977) theory (Echeverría Castro et al., 2020).

Accordingly, Usher and Pajares (2009) reviewed quantitative and qualitative measures and dissimilar measures of self-efficacy resources in mathematics. They found adjustment problems with the Bandura theory arising from similar procedures of building the test items or that, from a comprehensive viewpoint, they assessed overall self-efficacy. Thus they established their own scale for high school students, with a measurement model guaranteeing the test items addressed Bandura's (1977) four factors. They employed a three-stage methodology. The first stage involved a focus group, and the second and third stages were quantitative. This version was employed to perform some adjustments and

**Table 1.** Previous studies that adapted SMES

| Study | Sample | | | Grade | Country | n | Summary of goodness-of-fit statistics for SMES | | | | | Statistical technique used to verify factor structure | | VES |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Boy | Girl | Total | | | | $\chi^2$ (df) | CFI | SRMR | RMSEA | GFI | NFI | EFA | CFA | |
| Aziz and Azhar (2021) | 421 | 557 | 978 | 5 different high schools | Indonesia | 24 | 546.08 (246) | .910 | - | .070 | .908 | .90 | - | - | - |
| Chen (2010) | 973 | 1,063 | 2,036 | 7th-8th | Taiwan | 21 | 1,467.7 (183) | .965 | - | .058 | - | .096 | Principal components & varimax techniques | AMOS | - |
| Echeverría Castro et al. (2020) | 121 | 125 | 246 | University students | Mexico | 14 | 318.36 | .973 | - | .060 | - | - | Principal components & varimax techniques | AMOS | 67% |
| Kontas and Ozcan (2017) | 125 | 157 | 282 | 6th-7th-8th | Turkey | 24 | 552.79 (246) | .980 | .05 | .060 | - | - | - | AMOS | 43% |
| Usher and Pajares (2009) | 395 | 408 | 803 | 6th-7th-8th | USA | 24 | S-B$\chi^2$ (246)=601.21 | .960 | RMR=.04 | .040 | - | - | Maximum likelihood & ProMax oblique rotation | AMOS | - |
| Yurt (2014) | 408 | 342 | 750 | 6th-7th-8th | Turkey | 24 | 88.15 (233) | .950 | .07 | .070 | - | .90 | Unrotated principal component & equimax rotation | Maximum likelihood | 69% |

Note. n: Final numbers of items &VES: Variance explained by scale

execute exploratory factor analysis (EFA) with the estimate of ML and Promax oblique rotation. Ultimately, confirmatory factor analysis (CFA) of the model was executed after the instrument was employed once again to a different sample. The measurement model, which consists of 24 items and six items for each factor, was retained because the measurement model demonstrated a good fit (Echeverría Castro et al., 2020).

SMES was constructed by Usher and Pajares (2009) in the USA to use for middle school mathematics students. According to Usher and Pajares (2009), SMES has robust content validity, criterion validity, and internal consistency. Thus, SMES is psychometrically sound and can be reliably used to assess the mathematics self-efficacy of students in middle school. Corsi-Bunker (2015) stated that SMES is appropriate for middle schools, where students' ages are between 11 and 14. Moreover, SMES has been used in various studies (Campbell & Stanley, 2015; Chen, 2010; Gordon, 2012; Kandemir & Akbas-Perkmen, 2017), which illustrates that this scale is strong in measuring sources of mathematics self-efficacy. Moreover, according to Usher and Pajares (2009), "SMES can be adapted for use in other domains" (p. 89) and can, therefore, be used in different cultures or contexts. Thus, SMES has been translated and adapted to different contexts such as Indonesia, Mexico, Taiwan, and Turkey (see **Table 1**). SMES has better psychometric properties according to Kandemir and Akbas-Perkmen (2017), who recommended its use to the mathematics teachers in Turkey. **Table 1** shows some of the previous studies that have adapted SMES.

### Objectives

1. Translating SMES into Arabic.
2. Adapting and validating SMES for use in the field of education in Oman.
3. Providing evidence of the reliability and validity of SMES as a tool to measure the sources of self-efficacy of students in mathematics in Oman in light of Bandura's (1977) self-efficacy theory.

## METHOD

### Participants

This study sample comprises 700 eighth-grade students, that is 321 boys (46%) and 379 girls (54%), aged around 14 years. The study sample is drawn from second-cycle public schools of basic education in Al Batinah North Governorate, the Sultanate of Oman (the biggest governorate in Oman).

### Instruments

This study used three types of instruments:

(1) source of mathematics self-efficacy scale,

(2) mathematics motivational scale, and

(3) national mathematics test.

SMES developed by Usher and Pajares (2009). It has acceptable fit, with S-B v2(246)=601.21, p<.0001, comparative fit index (CFI)=.96, root mean square error approximation (RMSEA)=.04, and SRMR=.04. Moreover, all standardized factor loadings were significant at α=.05 level and ranged in magnitude from .61 to .83.

**Table 2.** Description of SMES

| | SMES: Four factors | | | | Total items of scale |
|---|---|---|---|---|---|
| | Mastery experience (ME) | Vicarious experience (VS) | Social persuasion (P) | Physiological state (PH) | |
| Number of items of each factor | 6 | 6 | 6 | 6 | 24 |
| Number of items belonging to each factor | 1-6 (SEF) | 7-12 (SEF) | 13-18 (SEF) | 19-24 (SEF) | |
| Reverse-scored items (negative) | ME-3 (SEF3) | - | - | All items | |
| Cronbach's alpha coefficients | .88 | .84 | .88 | .87 | |
| Intercorrelation | ME↔VS=.61<br>P↔PH=-.56 | VS↔P =.71<br>VS↔PH = -.45 | ME↔P=.83 | ME↔PH=-.70 | |

In addition, the response format on SMES allows individuals to rate themselves from one (definitely false) to six (definitely true). Moreover, the participants rate on a six-point Likert-type scale, which ranges from 1="not at all confident" to 6="absolutely confident". In addition, there were seven negatively worded statements in the scales ME-3 ( SEF3), PH-1 (SEF19),PH-2 (SEF20),PH-3 (SEF21),PH-4 (SEF22), PH-5(SEF23), and PH-6 (SEF24) with the others being positive (see **Table 2** for more details).

### Mathematics Motivational Scale

The researchers have used mathematics motivation scale (MMS) to measure students' motivation to learn mathematics. MMS is constructed by Hussien et al. (2012) in the UAE. Hussien et al. (2012) constructed MMS by using five theoretical subscales: external regulation, introjected regulation, intrinsic motivation, amotivation and identified regulation. MMS consists of 44 items, and a 4-point Likert scale ("1"=does not describe me at all; "2"=describes me a little; "3"=greatly describes me; "4"=describes me completely), which measures students' motivation to learn mathematics and relies on self-determination theory, where self-determination theory helps to understand human motivation (Zhang et al., 2016). Moreover, MMS was used by Omani scholars (Aljahwari & Aldhafri, 2017) and this gives more credibility to their results. Aljahwari and Aldhafri (2017) used MMS to measure the motivation of eighth-grade students and ran the MMS on Omani data before using the scale and found that internal consistency coefficients for intrinsic motivation ($\alpha$=0.92) and extrinsic motivation ($\alpha$=0.85) showed good reliability evidence.

### Mathematics Achievement

National mathematics achievement test for grade 8 was developed by the Omani Ministry of Education in the first semester of the academic year 2018-2019. An achievement test is described by Devi and Sharma (2013) as an exam of information or competence that relies on something studied or educated. Therefore, national mathematics achievement test is an assessment of knowledge or skill on mathematics-related material.

### Procedure

#### *First phase: Obtaining of permission*

The researchers first obtained permission from Usher and Pajares (2009) to translate SMES into Arabic. Then, the researchers went to the Ministry of Education to obtain permission to conduct a study in second-cycle public schools of basic education in the Al Batinah North Governorate. After being granted permission, the researchers chose six schools (boys' and girls' schools) as a randomly selected sample, as per the rules of sample selection. Subsequently, the researchers informed the selected schools and acquired permission from their administrators to use the research instrument on the eighth-grade students. Finally, with the help of the teachers, the researchers used the instruments of the study on the selected sample during the school day by distributing it to the students.

#### *Second phase: Translation of SMES into Arabic*

SMES was created in English; thus, the researchers translated it into Arabic using the forward-and-back translation approach (**Figure 1**). This method has been used by many researchers to translate scales and questionnaires (Erdvik et al., 2015; Sowtali et al., 2016; Sun & Pju, 2009). According to Gorecki et al. (2014), the translation procedure includes six steps. Accordingly, the first step involved forwarding the translation from English into Arabic. During this step, the researchers gave the scale to two translators for translation into Arabic. In the second step, the researchers, along with an English-language expert, reviewed and compared the two translations and came to a consensus on a definitive version. In the third step, the scale was translated backward from Arabic to English by two translators–not the ones who carried out the initial translation. The fourth step involved reviewing the backward and forward translations; the goal of this step, according to Gorecki et al. (2014), is to determine whether the equivalent versions affirm that satisfying procedures were followed before the pre-test. The fifth step focused on pre-testing the translated scale; the pre-test is performed to understand whether all items can be

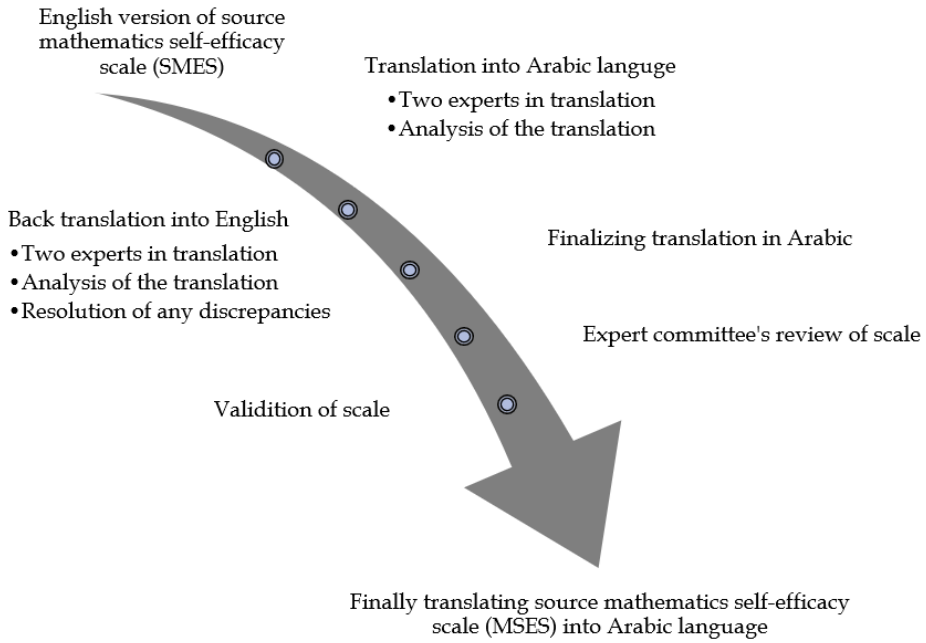


**Figure 1.** Diagram of translation process (Source: Authors' own elaboration)

understood and whether they are acceptable. It also aims to recognize and solve any existing issue with the translated questionnaire (e.g., to define whether each item and request is thorough and satisfactory). In this step, the researchers administered a "pre-test" on a sample (30-40) of students. The researchers selected the number of participants according to Arafat (2016), who said that pre-testing must involve at least 30-40 individuals.

Further, the researchers conducted a cognitive interview with six randomly chosen students; they were interviewed for around 30 minutes. The cognitive interview aimed to comprehend how the defendants would answer the survey items (i.e., the fundamental cognitive processes used in reading, understanding, interpretation questions, and formulating responses).

Davis et al. (2013) suggested asking the following questions in a cognitive interview:

> "What do you believe the question asked? Could you repeat the question in your own words? What came to their minds when they heard a specific phrase or term? Could you explain how you chose your response?" (p. 12).

The researchers used these questions in the cognitive interview to understand how the students thought. The sixth step concerned validation and reliability.

### *Third phase: Validation of SMES*

The validity of an instrument is defined as the capability of a tool to gauge the characteristics of the structure under research DeVon et al. (2007). Validity is of two types, namely

(a) translational validity and

(b) criterion validity, which are also known as construct validity.

Translational validity is further classified into two:

(a) content validity and

(b) face validity (Carroll, 2016; DeVon et al., 2007).

The researchers conducted the face and content validity tests of the instrument by sending the instrument of this study to experts. Most of the experts on the panel for testing the instrument were in the teaching fields of mathematics and curriculum design. The panel agreed with the researchers on most points of judgment and differed on a few. The panel provided oral as well as written comments. The researchers took the experts' perspectives and comments into consideration.

**Face validity:** Face validity refers to the extent to which an exam demonstrates the ability to gauge what it is intended to gauge (Volkmar, 2013). Lewis-Beck et al. (2003) defined face validity as an estimation of the extent to which a measure is obviously and explicitly for the structure it aims to evaluate. According to the literature review, the researchers of this study have identified a different method for evaluating face validity. Some researchers, such as Heale and Twycross (2015), viewed face validity as a subgroup of content validity, where specialists are questioned on their views regarding whether a tool assesses the concept it is intended to assess. Taherdoost (2016) evaluated the face validity of a questionnaire with regard to its viability, legibility, constancy of manner and structuring, and the simplicity of the dialect used.

Parsian and Dunning (2009) determined the face validity of a questionnaire by assessing each question in the questionnaire with regard to the intelligibility of the expression, the probability of the objective public being

**Table 3.** IFV of SMES

| No | Statement |
|---|---|
| 1 | Is the instruction of the instrument adequate? |
| 2 | Are instructions for respondents on first page proper? |
| 3 | Are instructions for respondents on the first page clear? |
| 4 | Are items of the instrument clear and unambiguous? |
| 5 | Is difficulty level of items proper for the respondents? |
| 6 | Is spelling of difficult words in the instrument correct? |
| 7 | Are items reasonable in relation to perceived purpose of instrument? |
| 8 | Is the print of the instrument legible? |

capable of responding to the queries, the design, and the manner. Oluwatayo (2012) stated that quantitative assessment of face validity could be performed by getting specialists in the relevant field of research to rate the appropriateness of the measuring tool for its intentional usage. Moreover, Oluwatayo (2012) provided the following criteria for assessing the face-to-face validity of experts:

1. The construction of the tool with regard to its structure and its well-thought form
2. The spacing of objects among lines
3. The adequacy of instruction on the tool
4. The legibility of print
5. No spelling errors

The researchers used the aforementioned criteria to evaluate the face validity of all instruments used in this study. Moreover, to rate the face validity of the criteria items, Taherdoost (2016) suggested using a dichotomous scale for a decisive choice of "yes" and "no", signaling a "favorable" and "unfavorable" item, respectively. Thus, the researchers used the criteria of face validity introduced by Oluwatayo (2012) and the dichotomous scale of Taherdoost (2016) to rate the items in the criteria.

With regard to SMES, the experts unanimously agreed that the face validity was good. **Table 3** explains index of face validity (IFV) of SMES; the researchers used the percentage of agreement statistic to measure the agreement among raters.

**Content validity:** In this stage, the researchers measured content validity. Salkind (2010) defined content validity as the degree to which the clauses on an exam are equally representative of the entire field of the exam to be measured. Content validity, according to Oluwatayo (2012) defined content validity as a hypothetical idea that concentrates on degree to which measuring tool reveals proof of equality and thorough covering of the field of clauses that it claims to cover.

This stage has two steps: First, the researchers used the content validity index, providing proof of content validity by calculating a content validity indicator according to the instructions of Polit and Beck (2006). To achieve this, the researchers designed a content validity index for the panel members by rating the tool clauses with regard to the lucidity and relevance of the construct underlying the research as per the theoretical introductions of the construct itself and its aspects (Zamanzadeh et al., 2015).

Second, the content validity of the tool is defined using the opinions of the committee of specialists (Zamanzadeh et al. ,2015). The number of specialists should at least be five. However, Opara and Magnus-Arewa (2017) used two experts to judge the content validity and face validity. Therefore, the researchers chose two experts among lecturers at Malaysian and Omani universities.

Furthermore, Drost (2011) suggested two ways to assess content validity:

(a) asking several questions regarding the tool or the exam and

(b) questioning the point of view of the specialist judges in the field.

Thus, the researchers chose to question the points of view of specialists or committee judges in this field. According to Zamanzadeh et al. (2015), the committee members are asked to assess the tool clauses with regard to

(a) clarity and

(b) its relevance to the construct fundamental the study as per the theoretical definitions of the construct itself and its aspects on a four-point ordinal scale–(1) not relevant, (2) somewhat relevant, (3) quite relevant, and (4) highly relevant. It was the same for clarity–(1) not clear, (2) somewhat clear, (3) quite clear, and (4) highly clear.

Sowtali et al. (2016) determined the criterion of translational content validity by assessing each clause across four aspects utilizing a dichotomous response scale: "clear=1" and "not clear=0." The four dimensions were

(a) consistency of the item regarding a content area,

(b) clarity of the item's wording,

(c) perceived item difficulty, and

(d) the item's potential for being contained in a revised version of the test.

Sowtali et al. (2016) stated the same criteria for content validity. This study utilized two designs, one of face validity and the other of content validity, to assess the face and content validity of all instruments of this study. The researchers sent SMES to the experts chosen to judge the scale on its face and content validity.

The rate of consistency among the raters for SMES (content validity) was 0.875; hence, the agreement between inter-raters was moderate. In terms of clarity and difficulty, it was 1.00; this indicates that the agreement between inter-raters was almost perfect, according to McHugh (2012).

**Construct validity:** Construct validity is defined as the extent to which the measures used, usually surveys, essentially assess the hypothesis or theory they are measuring (Gellman & Turner, 2013). For the purpose of analysis, the researchers performed EFA and CFA. EFA helps the researchers make indications based on the internal structure by keeping only those items with properly high loadings on factor 1 for competence, the construct of concern (Tavakol & Wetzel, 2020). Suhr (2006) defined CFA as a statistical method used to validate the factor construction of a group of monitored variables. Said et al. (2011) stated that CFA is used to test the validity and reliability of research instruments. Moreover, DiStefano and Hess (2005) stated that CFA gives proof of construct validity in psychological assessments. Previously, studies have used EFA and CFA with diverse types of extraction methods and rotation to explore the factorial structure of the scale (Yurt, 2014) (**Table 1**). However, in the original study, Usher and Pajares (2009) used EFA and CFA by SPSS and AMOS, respectively, besides maximum likelihood (ML) and Promax oblique rotation during EFA.

Accordingly, in the current study, the researchers used the same analytical techniques used in the previous studies for validating the scale (see **Table 1**). The researchers ran an EFA to determine the nature of factors influencing a set of variables (Mvududu & Sink, 2013), with a principal component analysis and varimax rotation by SPSS. This is because the assumption of principal components is that the original factors are orthogonal and not correlated with each other, which is not accurate for most social sciences.

Other extraction and rotation techniques that permit the extracted factors to correlate with each other have also been tried by researchers, but they produced unaccountable mathematical solutions (Chen, 2010). Therefore, EFA and CFA were used, as was done in the original study, by utilizing the scores of Omani students to explore the factorial structure of the scale. Thus, the researchers were able to verify the construct validity of SMES.

## RESULTS

The outcomes of the current study have been presented in two sections: EFA and CFA.

### EFA

EFA has two assumptions that must be considered initially before performing the analysis. The first assumption is sampling adequacy, tested using the Kaiser-Meyer-Olkin coefficient, which was .938, meaning there was sampling adequacy. The second, the normal distribution of the data, was tested using the Bartlett test; the results were $\chi^2$=7,676.156, with p<0.001 (Yurt, 2014). Both met the assumption criteria of the test.

**Table 4.** Rotated matrix of self-efficacy instrument at the end of EFA

| | Factors | | | |
|---|---|---|---|---|
| | 1 | 2 | 3 | 4 |
| SEF17 | .799 | | | |
| SEF15 | .732 | | | |
| SEF14 | .731 | | | |
| SEF16 | .728 | | | |
| SEF18 | .698 | | | |
| SEF13 | .631 | | | |
| SEF24 | | .779 | | |
| SEF21 | | .779 | | |
| SEF22 | | .739 | | |
| SEF19 | | .732 | | |
| SEF20 | | .726 | | |
| SEF23 | | .712 | | |
| SEF9 | | | .804 | |
| SEF7 | | | .728 | |
| SEF10 | | | .704 | |
| SEF8 | | | .667 | |
| SEF11 | | | .480 | |
| SEF12 | | | .453 | |
| SEF2 | | | | .801 |
| SEF6 | | | | .674 |
| SEF5 | | | | .627 |
| SEF1 | | | | .598 |
| SEF4 | | | | .527 |

The results of EFA show that the present outcomes are the same results as in the original study in terms of the number of factors. Without restrictions on the number of factors, a four-factor structure is observed, similar to that found in the scale by Usher and Pajares (2009); for factors in the extraction process (see **Table 1**). Moreover, these results are consistent with the theory of self-efficacy of Bandura (1977). However, item ME-3, which stated that "even when I study very hard, I do poorly in math", was deleted, because this item was essentially a factor of mastery experience, but the analysis shows that ME-3 is actually a factor of arousal and physiological states. Thus, this item had caused an overlap between these two factors. This happened because the students interpreted item ME-3 as being related to their arousal and physiological state. This result could be attributed to the translation from one language to another one and the different cultural background of the students, which led to the difference in understanding item ME-3. Therefore, the researchers decided to remove item ME-3 because it carried a different meaning for the respondents. Moreover, item ME-3 cannot be considered under a different factor; indeed, this removal was regarded more appropriate, as it would not impact the final result of the self-efficacy of the students. Therefore, 23 items remained, as shown in **Table 4**. The model explained 60.151 % of the variance. Factor 1 accounted for 38.177% of the variance, factor 2 for 10.490%, factor 3 for 6.498%, and factor 4 for 4.986%. The results of the factor analysis are presented in **Table 4**.

**Table 5.** Means, standard deviations, & correlations for final sources of self-efficacy items (n=700) (item-total correlations between each item & its subscale counterparts appear diagonally & items within each given subscale appear in **bold**)

| Item | *M* | *SD* | SEF1 | SEF2 | SEF4 | SEF5 | SEF6 | SEF7 | SEF8 | SEF9 | SEF10 | SEF11 | SEF12 | SEF13 | SEF14 | SEF15 | SEF16 | SEF17 | SEF18 | SEF19 | SEF20 | SEF21 | SEF22 | SEF23 | SEF24 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| SEF1 | 4.591 | 1.391 | **0.653** | | | | | | | | | | | | | | | | | | | | | | |
| SEF2 | 5.151 | 1.272 | 0.54 | **0.617** | | | | | | | | | | | | | | | | | | | | | |
| SEF4 | 4.633 | 1.423 | 0.574 | 0.43 | **0.558** | | | | | | | | | | | | | | | | | | | | |
| SEF5 | 5.147 | 1.16 | 0.397 | 0.432 | 0.337 | **0.547** | | | | | | | | | | | | | | | | | | | |
| SEF6 | 4.586 | 1.39 | 0.46 | 0.487 | 0.383 | 0.545 | **0.603** | | | | | | | | | | | | | | | | | | |
| SEF7 | 4.916 | 1.411 | 0.338 | 0.228 | 0.249 | 0.265 | 0.378 | **0.593** | | | | | | | | | | | | | | | | | |
| SEF8 | 4.683 | 1.479 | 0.363 | 312 | 0.285 | 0.307 | 0.346 | 0.435 | **0.59** | | | | | | | | | | | | | | | | |
| SEF9 | 4.734 | 1.538 | 0.314 | 0.064 | 0.337 | 0.271 | 0.301 | 0.548 | 0.45 | **0.618** | | | | | | | | | | | | | | | |
| SEF10 | 4.754 | 1.382 | 0.355 | 0.251 | 0.267 | 0.298 | 0.351 | 0.394 | 0.549 | 0.479 | **0.612** | | | | | | | | | | | | | | |
| SEF11 | 4.66 | 1.407 | 0.463 | 0.375 | 0.375 | 0.357 | 0.506 | 0.394 | 0.399 | 0.399 | 0.451 | **0.593** | | | | | | | | | | | | | |
| SEF12 | 4.826 | 1.402 | 0.494 | 0.403 | 0.393 | 0.364 | 0.407 | 0.437 | 0.378 | 0.424 | 0.402 | 0.582 | **0.592** | | | | | | | | | | | | |
| SEF13 | 4.587 | 1.579 | 0.483 | 0.397 | 0.439 | 0.379 | 0.41 | 0.315 | 0.373 | 0.322 | 0.349 | 0.472 | 0.503 | **0.667** | | | | | | | | | | | |
| SEF14 | 4.119 | 1.685 | 0.482 | 0.388 | 0.407 | 0.401 | 0.404 | 0.299 | 0.327 | 0.315 | 0.331 | 0.462 | 0.514 | 0.574 | **0.733** | | | | | | | | | | |
| SEF15 | 4.26 | 1.694 | 0.471 | 0.35 | 0.405 | 0.393 | 0.378 | 0.324 | 0.358 | 0.312 | 0.313 | 0.438 | 0.441 | 0.566 | 0.633 | **0.718** | | | | | | | | | |
| SEF16 | 4.031 | 1.671 | 0.428 | 0.355 | 0.397 | 0.39 | 0.401 | 0.278 | 0.339 | 0.305 | 0.338 | 0.419 | 0.428 | 0.514 | 0.611 | 0.558 | **0.696** | | | | | | | | |
| SEF17 | 3.864 | 1.728 | 0.424 | 0.353 | 0.37 | 0.336 | 0.332 | 0.236 | 0.301 | 0.251 | 0.273 | 0.397 | 0.432 | 0.525 | 0.565 | 0.597 | 0.573 | **0.709** | | | | | | | |
| SEF18 | 4.186 | 1.599 | 0.483 | 0.34 | 0.392 | 0.353 | 0.381 | 0.31 | 0.358 | 0.317 | 0.37 | 0.439 | 0.448 | 0.533 | 0.543 | 0.533 | 0.553 | 0.59 | **0.679** | | | | | | |
| SEF19 | 2.547 | 1.774 | -0.321 | -0.312 | -0.224 | -0.267 | -0.241 | -0.185 | -0.203 | -0.09 | -0.182 | -0.261 | -0.282 | -0.306 | -0.293 | -0.304 | -0.238 | -0.211 | -0.231 | **0.587** | | | | | |
| SEF20 | 3.013 | 1.768 | -0.189 | -0.136 | -0.158 | -0.138 | -0.184 | -0.115 | -0.173 | -0.072 | -0.164 | -0.188 | -0.201 | -0.186 | -0.216 | -0.237 | -0.218 | -0.186 | -0.212 | 0.463 | **0.676** | | | | |
| SEF21 | 2.921 | 1.819 | -0.277 | -0.196 | -0.218 | -0.228 | -0.247 | -0.203 | -0.204 | -0.127 | -0.157 | -0.258 | -0.257 | -0.244 | -0.243 | -0.258 | -0.216 | -0.195 | -0.22 | 0.527 | 0.531 | **0.679** | | | |
| SEF22 | 2.76 | 1.83 | -0.361 | -0.258 | -0.295 | -0.259 | -0.285 | -0.187 | -0.168 | -0.125 | -0.183 | -0.31 | -0.339 | -0.317 | -0.32 | -0.315 | -0.308 | -0.255 | -0.302 | 0.492 | 0.495 | 0.496 | **0.676** | | |
| SEF23 | 2.644 | 1.799 | -0.397 | -0.266 | -0.288 | -0.286 | -0.328 | -0.259 | -0.214 | -0.239 | -0.271 | -0.416 | -0.37 | -0.351 | -0.369 | -0.371 | -0.297 | -0.3 | -0.341 | 0.527 | 0.416 | 0.513 | 0.589 | **0.718** | |
| SEF24 | 2.78 | 1.832 | -0.354 | -0.281 | -0.312 | -0.225 | -0.266 | -0.172 | -0.204 | -0.194 | -0.217 | -0.331 | -0.329 | -0.316 | -0.297 | -0.274 | -0.247 | -0.226 | -0.3 | 0.568 | 0.443 | 0.585 | 0.588 | 0.603 | **0.654** |

**Table 5** shows the correlation matrix and item-total correlations for the dependent (examined) variables in the model. The inter-item correlations among the six items designed to measure three sources (vicarious experience, social persuasions, and physiological state) and five items designed to measure mastery experience ranged from -.41 to 0.63. The items in all of the four subscales demonstrated adequate internal consistency, with Cronbach's alpha coefficients above (0.81), which is considered robust (Taber, 2018), as was in original study. Specifically, Cronbach's alpha coefficient was 0.80 for mastery experience, 0.83 for vicarious experience, 0.88 for social persuasions, and 0.86 for physiological state.

## CFA

Most previous studies that applied CFA to SMES to verify the factor structure of a set of noted variables used the following goodness of fit indices: root mean square error of approximation (RMSEA), standardized root mean square residual (SRMR), CFI, and goodness of fit index (GFI) (see **Table 1**). The current study used the same goodness of fit index that was used by previous studies and, specifically, the original study: RMSEA, SRMR, CFI, $\chi^2$, and df. The indices of convergent and divergent validity for this scale were all significant ($p<0.001$).

The model shown in **Figure 2** demonstrated adequate fit, with $\chi^2=767.8$, df=224, p<.0001, CFI=.928, SRMR=.0478, and RMSEA=.059. The standardized factor loadings in the model were significant at the α=.001 level and ranged in magnitude from .61 to .79 (see **Figure 2**). The four sources' factors demonstrated intercorrelations ranging in magnitude from -.48 (between vicarious experience and physiological state) to .78 (between social persuasions and mastery experience). There was a potent correlation between the mastery experience and social persuasions items, as was in the original study (Usher & Pajares, 2009) (see **Figure 2**).

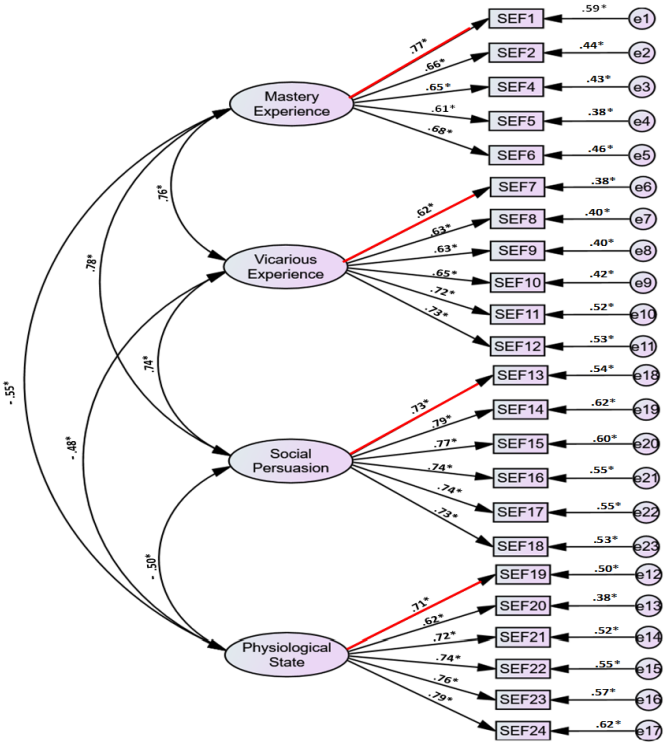


**Figure 2.** Measurement model of SMES ($\chi^2=767.8$, df=224, CFI=.928, SRMR=.0478, & RMSEA=.059) (p< 0.001) (Source: Authors' own elaboration)

The researchers formulated a second-order model that demonstrated the convergence between the first-order and second-order factors as well as the existence of a discriminant between them, considering that they have moderate correlations (see **Figure 3**). This is consistent with Echeverría Castro et al. (2020).

**Table 6** provides information regarding the convergence between the first-order and second-order factors for both boys and girls. Also, it shows the existence of a discriminant between the first-order and second-order factors for both boys and girls considering that they have moderate correlations.

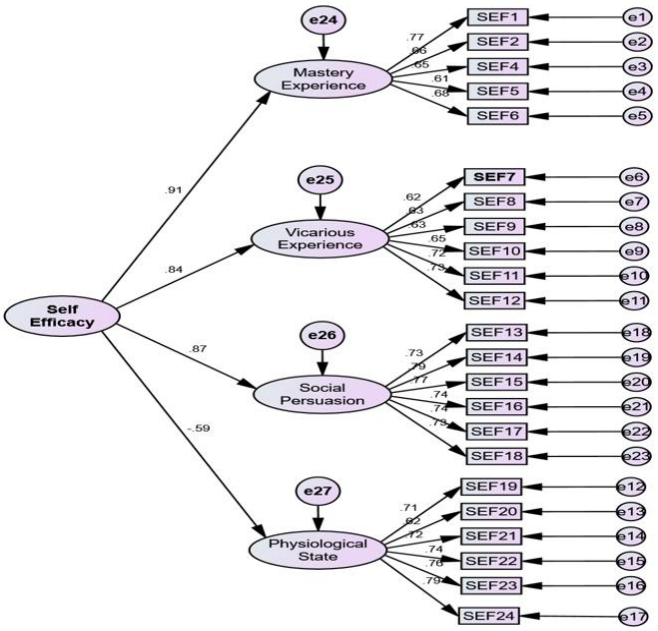


**Figure 3.** Measurement model for conformation of second-order factors of mathematics self-efficacy ($\chi^2$=767.8, df=224, CFI=.928, SRMR=.0478, & RMSEA=.059) (p< 0.001) (Source: Authors' own elaboration)

**Table 6.** Correlations with second-order factors of mathematics self-efficacy

| Construct | Boys | Girls |
|---|---|---|
| Mastery experience<---Self-efficacy | 0.865* | 0.939* |
| Vicarious experience<---Self-efficacy | 0.869* | 0.844* |
| Social persuasion<---Self-efficacy | 0.88* | 0.858* |
| Physiological state<---Self-efficacy | -0.482* | -0.666* |

Note. *p<.001

**Tests for measuring invariance:** **Table 7** provides information regarding the factorial weight for both boys and girls, which was above 0.5. The researchers conducted CFAs for the measurement models of boys and girls. The analysis points to models' levels of fit being acceptable. We identified an almost identical model for boys and girls when testing for factorial invariance (see **Table 8**). The measurement model was invariant for both genders, with $\Delta\chi^2$=30.2, $\Delta$df=23, and p>.05 (see **Table 9**). It is evident from the nonsignificant chi square statistic that the null hypothesis cannot be rejected, which states that the model proposed differs from the population model (Usher & Pajares, 2009). These results are consistent with (Echeverría Castro et al., 2020; Usher & Pajares, 2009).

**Proof of construct validity:** The researchers evaluated the convergent validity by examining the correlations among mathematics achievement, intrinsic motivation for mathematics Hussien et al. (2012), and the four sources of self-efficacy. As can be seen from **Table 10**, the correlations between the sources and mathematics achievement and mathematics intrinsic motivation for mathematics were all statistically significant (p<.001). Moreover, the researchers found a statistically significant relationship between self-efficacy and the sources of self-efficacy at (p<.001), ranging from an absolute value of .34 to .82, consistent with the results of Usher and Pajares (2009). Mathematics achievement and mathematics intrinsic motivation for mathematics were correlated with the four sources of self-efficacy, which offers compelling evidence for the criterion validity of the source's subscales. According to the aforementioned information, the scale is working well.

**Table 7.** CFA in both samples (factorial weight)

| Item | | Girls | Boys |
|---|---|---|---|
| SEF1 | Mastery experience | 0.765 | 0.754 |
| SEF2 | | 0.677 | 0.658 |
| SEF4 | | 0.646 | 0.683 |
| SEF5 | | 0.631 | 0.605 |
| SEF6 | | 0.693 | 0.670 |
| SEF7 | Vicarious experience | 0.652 | 0.571 |
| SEF8 | | 0.597 | 0.626 |
| SEF9 | | 0.571 | 0.637 |
| SEF10 | | 0.639 | 0.621 |
| SEF11 | | 0.697 | 0.727 |
| SEF12 | | 0.727 | 0.720 |
| SEF19 | Physiological state | 0.716 | 0.698 |
| SEF20 | | 0.595 | 0.616 |
| SEF21 | | 0.692 | 0.726 |
| SEF22 | | 0.721 | 0.739 |
| SEF23 | | 0.781 | 0.734 |
| SEF24 | | 0.788 | 0.785 |
| SEF13 | Social persuasion | 0.738 | 0.728 |
| SEF14 | | 0.817 | 0.782 |
| SEF15 | | 0.788 | 0.757 |
| SEF16 | | 0.800 | 0.728 |
| SEF17 | | 0.765 | 0.736 |
| SEF18 | | 0.758 | 0.709 |

**Table 8.** Summary of goodness-of-fit statistics for final sources of self-efficacy measurement model by subgroup

| Subgroup model | $\chi^2$ | df | CFI | SRMR | RMSEA |
|---|---|---|---|---|---|
| Girls | 667.9 | 224 | . 90 | .056 | .072 |
| Boys | 443.6 | 224 | .93 | .051 | .055 |

Note. p<.001

## DISCUSSION

This study sought to translate, adapt, and validate SMES developed by Usher and Pajares (2009) into Arabic in light of the self-efficacy theory of Bandura (1977), providing evidence of the reliability and validity of SMES as a tool to measure the sources of self-efficacy in

**Table 9.** Tests for invariance of final sources of self-efficacy measurement model across gender

| Subgroup model | $\chi^2$ | df | CFI | SRMR | RMSEA | RMSEA 90% CI | Model compassion | $\Delta\chi^2$ | $\Delta$df | $\Delta$CFI |
|---|---|---|---|---|---|---|---|---|---|---|
| Model 1 | 1,124.0 | 454 | .91 | .059 | .046 | .043, 0.49 | - | - | - | - |
| Model 2 | 1,154.2 | 477 | .91 | .063 | .045 | .042,048 | - | 30.2 | 23 | .01 |

Note. Model 1: Gender: Model 1: Configural (no constraints) & Model 2: Gender: Model 2: Factor loadings invariant

**Table 10.** Means, standard deviations, & zero-order correlations for variables for sample (n=374)

| Variable | M | SD | TEST | MS | P | VS | PH | IM |
|---|---|---|---|---|---|---|---|---|
| 1. TEST | 19.2400 | 11.25000 | - | | | | | |
| 2. MS | 23.8984 | 5.37947 | 0.548*** | - | | | | |
| 3. P | 27.5722 | 6.81536 | 0.434*** | 0.640*** | - | | | |
| 4. VS | 24.6979 | 8.30370 | 0.481*** | 0.667*** | 0.622*** | - | | |
| 5. PH | 16.3529 | 8.59069 | -0.439*** | -0.429*** | -0.294*** | -0.380*** | - | |
| 6. IM | 49.4064 | 11.23145 | 0.281*** | 0.473*** | 0.533*** | 0.531*** | -0.334*** | - |

Note. MS: Mastery experience; P: Social persuasion; VS: Vicarious experience; PH: Physiological state; IM: Intrinsic motivation; TEST: Mathematics achievement test; *p<.05; **p<.01; & ***p<.001

mathematics of Omani students. The researchers, to achieve the core aims of the current study, have translated and validated the scale by conducting EFA and CFA.

The EFA reveals that SMES has four (factors) sources as in the original scale, which concurs with the earlier studies that have adapted the scale (Kontas & Ozcan, 2017; Usher & Pajares, 2009; Yurt, 2014). However, it also demonstrates that item ME-3, which states that "even when I study very hard, I do poorly in math", is inconsistent with the other items or factors in the scale. This might be because this item carries a different meaning for the respondents (due to cultural differences). Therefore, the researchers dropped item ME-3 from factor mastery experience, and the number of items in SMES reduced to 23. This result is in line with those of Chen (2010), whose study was conducted in Taiwan. Chen (2010) deleted item ME-3 when adapting SMES to Taiwanese culture. Moreover, the item has a low item-total correlation, compared to the other items on the subscale, and it was the only item on the subscale that raised the reliability alpha when dropped. Thus, Chen deleted item ME-3 from the scale.

Therefore, ultimately, SMES has 23 items, reflecting the four sources of mathematics self-efficacy. Moreover, the scale has adequate reliability (Cronbach's alpha α). Cronbach's alpha coefficient was above 0.81, indicating the scale's robustness (Taber, 2018), in line with the results found by (Usher & Pajares, 2009). Moreover, Cronbach's alpha coefficients were .80 for mastery experience, 0.83 for vicarious experience, .88 for social persuasions, and 0.86 for the physiological state. This result gives strong evidence regarding the internal consistency of the scale.

CFA shows that the 23-item SMES has good fit, where $\chi^2(224)=767.8$, CFI=.928, SRMR=.0478, and RMSEA=.059. The results of CFA are consistent with the past studies adapting SMES (Kontas & Ozcan, 2017; Usher & Pajares, 2009; Yurt, 2014). Regarding the convergent validity, the researchers ran a second-order model, which confirmed convergence between the first-order and the second-order factors of mathematics self-efficacy as well as the existence of a discriminant between them, considering that they have moderate correlations (see **Figure 2** and **Table 9**). This result provides assurance that the scale has a good psychometric property similar to the original version. The factorial weights are significant and above 0.5 for all boys and girls.

The tests for measuring invariance with the gender construct variable shows an acceptable fit for boys and girls. The measurement model was invariant for both genders, with $\Delta\chi^2=30.2$ and df=23 (see **Table 9**). The insignificant chi-square statistic provides proof regarding the null hypothesis cannot be rejected, which says the model assumed does not vary from the population model (Usher & Pajares, 2009). Thus, the configurational invariance shows that the basic measurement models are equal, and the factorial weights are comparable for boys and girls.

The convergent validity of SMES was backed by the strong correlation relationship between the four sources of self-efficacy and mathematics achievement and mathematics intrinsic motivation for mathematics, which offers proof regarding the scale's convergent validity. Regarding the construct validity, the results of zero-order correlations for the variables showed that they were correlated with the four sources of self-efficacy and mathematics achievement and mathematics intrinsic motivation for mathematics (see **Table 7**), which also proves the scale's construct validity. This presents persuasive evidence for the criterion validity of the source's subscales.

As a result of the results, there is an acceptable level of structure validity in SMES. SMES also appears to have acceptable internal consistency. Furthermore, the psychometrical properties of SMES are proximate to the original sources of middle school mathematics scale, which was constructed by Usher and Pajares (2009).

## Implications

The findings of this study are substantial for educators and scholars. The findings have shown that the sources of the mathematics self-efficacy scale are psychometrically sound and can be dependably used to gauge the mathematics self-efficacy of students in middle school. And this will help educators improve the self-efficacy beliefs of their students. Furthermore, this study opened the door for more research on self-efficacy and its sources in Arab culture (Oman).

### Contributions

1. This study provided an instrument to measure students' sources of mathematics self-efficacy in middle school in Oman and in Arab culture by translating sources of mathematics self-efficacy scale, into Arabic in light of Bandura's (1977) self-efficacy theory.
2. This study examined the validity of the items of SMES across different contexts as one of the recommendations of Usher and Pajares (2009). Whereas this study found the same results as that found by Usher and Pajares (2009).

## CONCLUSIONS

The process of translation, adaption, and validation of SMES into Arabic in light of Bandura's (1977) theory of self-efficacy has demonstrated the reliability and validity of SMES as a tool to measure sources of self-efficacy of students in mathematics in the context of Oman. The Omani SMES consists of 23 items, representing four sources: mastery experience, social persuasion, vicarious experience, and physiological state. According to the aforementioned, SMES is psychometrically sound and can be dependably used to gauge the mathematics self-efficacy of students in middle school in Oman.

**Author contributions:** All authors have sufficiently contributed to the study and agreed with the results and conclusions.

**Funding:** No funding source is reported for this study.

**Acknowledgments:** The authors would like to thank those who supported research process.

**Ethical statement:** Authors stated that the study was approved by the Human Resources Development Department Vocational Training Development Center (Ministry of Education, Oman) on 17 September 2018 with Approval Code: 28190380.

**Declaration of interest:** No conflict of interest is declared by authors.

**Data sharing statement:** Data supporting the findings and conclusions are available upon request from the corresponding author.

## REFERENCES

Aljahwari, F. S., & Aldhafri, S. S. (2017). Relationship of academic self-efficacy with psychological adjustment among students in grades 7-12 in the Sultanate of Oman. *Journal of Educational and Psychological Studies, 12*(1), 163-178. https://journals.squ.edu.om/index.php/jeps/article/viewFile/2426/2323

Arafat, S. (2016). Cross cultural adaptation & psychometric validation of instruments: Step-wise description. *International Journal of Psychiatry, 1*(4), 1-4. https://doi.org/10.33140/IJP/01/01/00001

Artino, A. R. (2012). Academic self-efficacy: From educational theory to instructional practice. *Perspectives on Medical Education, 1*(2), 76-85. https://doi.org/10.1007/S40037-012-0012-5

Ayotola, A., & Adedeji, T. (2009). The relationship between mathematics self-efficacy and achievement in mathematics. *Procedia-Social and Behavioral Sciences, 1*(1), 953-957. https://doi.org/10.1016/j.sbspro.2009.01.169

Aziz, T. A., & Azhar, E. (2021). The validity and reliability study of the Indonesian version of the sources of mathematics self-efficacy scale. *AIP Conference Proceedings, 2331*, 020038. https://doi.org/10.1063/5.0041691

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*(2), 191-215. https://doi.org/10.1037/0033-295X.84.2.191

Bandura, A. (1994). Self-efficacy. In V. S. Ramachaudran (Ed.), Encyclopedia of human behavior (pp. 71-81). https://motivation.uky.edu/self-efficacy-overview/

Bandura, A. (2012). On the functional properties of perceived self-efficacy revisited. *Journal of Management, 38*(1), 9-43. https://doi.org/10.1177/0149206311410606

Britner, S. L., & Pajares, F. (2006). Sources of science self-efficacy beliefs of middle school students. *Journal of Research in Science Teaching, 43*(5), 485-499. https://doi.org/10.1002/tea.20131

Campbell, D. T., & Stanley, J. C. (2015). *Experimental and quasi-experimental designs for research*. https://www.sfu.ca/~palys/Campbell&Stanley-1959-Exptl&QuasiExptlDesignsForResearch.pdf

Carroll, S. R. (2016). *How to successfully deal with your dissertation data*. BookBaby.

Chang, S. H. (2006). *Testing the self-care self-efficacy enhancement program aimed at improving BADL performance for Chinese nursing home elders* [Doctoral dissertation, University of Arizona]. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.601.1421&rep=rep1&type=pdf

Chen, Y.-C. (2010). *Sources of mathematics self-efficacy and predictors of mathematics achievement among seventh- and eighth-grade Taiwanese students* [Doctoral dissertation, University of Kentucky]. https://search.proquest.com/docview/919564795?pq-origsite=gscholar

Corsi-Bunker, A. (2015). Guide to the education system in the United States. *University of Minnesota*. https://isss.umn.edu/publications/USEducation/2.pdf

Davis, E., Davern, M., & Waters, E. (2013). *CPQOL cerebral palsy quality of life questionnaire for adolescents (CP QOL-Teen) manual*. https://www.ausacpdm.org.au/wp-content/uploads/2019/01/CPQOL-Teen-manual-V3.pdf

Devi, S., & Sharma, H. (2013). Construction of an achievement test for the students of VIII class in the subject of mathematics. *International Journal of Scientific Research, 2*(7), 41-43. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.680.8243&rep=rep1&type=pdf

DeVon, H. A., Block, M. E., Moyle-Wright, P., Ernst, D. M., Hayden, S. J., Lazzara, D. J., Savoy, S. M., & Kostas-Polston, E. (2007). A psychometric toolbox for testing validity and reliability. *Journal of Nursing Scholarship, 39*(2), 155-164. https://doi.org/10.1111/j.1547-5069.2007.00161.x

DiStefano, C., & Hess, B. (2005). Using confirmatory factor analysis for construct validation: An empirical review. *Journal of Psychoeducational Assessment, 23*(3), 225-241. https://doi.org/10.1177/073428290502300303

Drost, E. A. (2011). Validity and reliability in social science research. *Education Research and Perspectives, 38*(1), 105. https://pdfs.semanticscholar.org/0815/34bfe6cf8dd0da1f40704098366f368da3e8.pdf

Echeverría Castro, S. B., Sotelo Castillo, M. A., Acosta Quiroz, C. O., & Barrera Hernández, L. F. (2020). Measurement model and adaptation of a self-efficacy scale for mathematics in university students. *Sage Open, 10*, 1. https://doi.org/10.1177/2158244019899089

Erdvik, I. B., Øverby, N. C., & Haugen, T. (2015). Translating, reliability testing, and validating a Norwegian questionnaire to assess adolescents' intentions to be physically active after high school graduation. *Sage Open, 5*(2), 1-6. https://doi.org/10.1177/2158244015580374

Evans, J. A. (2015). *Gender, self-efficacy, and mathematics achievement: An analysis of fourth grade and eighth grade TIMSS data from the United States* [Doctoral dissertation, Lesley University]. https://digitalcommons.lesley.edu/cgi/viewcontent.cgi?article=1080&context=education_dissertations

Gao, J. (2019). Sources of mathematics self-efficacy in Chinese students: A mixed-method study with q-sorting procedure. *International Journal of Science and Mathematics Education, 18*(4), 713-732. https://doi.org/10.1007/s10763-019-09984-1

Gellman, M. D., & Turner, J. R. (2013). *Encyclopedia of behavioral medicine*. Springer. https://doi.org/10.1007/978-94-007-6007-3

Gordon, D. (2012). *The relationship between self-efficacy, help-seeking behaviors and student achievement among middle level mathematics students in an interactive learning environment* [Doctoral dissertation, TUI University]. https://search.proquest.com/docview/928458598?pq-origsite=gscholar

Gorecki, C., Brown, J. M., Briggs, M., Coleman, S., Dealey, C., McGinnis, E., Nelson, E. A., Stubbs, N., Wilson, L., & Nixon, J. (2014). *Language translation & cross-cultural adaptation guideline*. https://ctru.leeds.ac.uk/wp-content/uploads/2019/01/Translation_Guidelines_10JAN14_finalv1.0.pdf

Heale, R., & Twycross, A. (2015). Validity and reliability in quantitative studies. *Evidence-Based Nursing, 18*(3), 66-67. https://doi.org/10.1136/eb-2015-102129

Hendricks, K. S. (2016). The sources of self-efficacy: Educational research and implications for music. *National Association for Music Education, 35*(1), 32-38. https://doi.org/10.1177/8755123315576535

Husain, U. K. (2014, December). *Relationship between self-efficacy and academic motivation* [Paper presentation]. The 2014 International Conference on Economics, Education and Humanities (ICEEH'14), Indonesia.

Hussien, J. H., Alsawaie, O., Alsartawi, A., Alghazo, I., & Tibi, S. (2012). Developing mathematics motivation scale for the United Arab Emirates. *Journal of Educational and Psychological Studies, 6*(3), 1-11. https://doi.org/10.12816/0002538

Kandemir, M. A., & Akbas-Perkmen, R. (2017). Examining validity of sources of mathematics self-efficacy scale in Turkey. *European Journal of Education Studies, 3*(33), 69-87. https://files.eric.ed.gov/fulltext/ED580955.pdf

Kaya, D., & Bozdag, H. C. (2016). Resources of mathematics self-efficacy and perception of science self-efficacy as predictors of academic achievement. *European Journal of Contemporary Education, 18*(4), 438-451. https://doi.org/10.13187/ejced.2016.18.438

Kehoe, M. (2013). *Make that grade organisational behaviour*. Gill & Macmillan Ltd.

Kontas, H., & Ozcan, B. (2017). Adapting sources of middle school mathematics self-efficacy scale to Turkish culture. *International Journal of Evaluation and Research in Education, 6*(4), 288-294. https://doi.org/10.11591/ijere.v6i4.10771

Kumar, R., & Lal, R. (2006). The role of self-efficacy and gender difference among the adolescents. *Journal of the Indian Academy of Applied Psychology, 32*(3), 345-350. https://psycnet.apa.org/record/2006-11517-013

Lewis-Beck, M., Bryman, A. E., & Liao, T. F. (2003). *The SAGE encyclopedia of social science research methods*. SAGE. https://doi.org/10.4135/9781412950589

Liu, X., & Koirala, H. (2009, October). *The effect of mathematics self-efficacy on mathematics achievement of high school students* [Paper presentation]. The 2009 Annual Conference of the Northeastern Educational Research Association, University of Connecticut, Connecticut, USA.

Margolis, H., & McCabe, P. P. (2006). Improving self-efficacy and motivation: What to do, what to say. *Intervention in School and Clinic, 41*(4), 218-227. https://doi.org/10.1177/10534512060410040401

McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica, 22*(3), 276-282. https://doi.org/10.11613/BM.2012.031

Meral, M., Colak, E., & Zereyak, E. (2012). The relationship between self-efficacy and academic performance. *Procedia-Social and Behavioral Sciences, 46*, 1143-1146. https://doi.org/10.1016/j.sbspro.2012.05.264

Mvududu, N. H., & Sink, C. A. (2013). Factor analysis in counseling research and practice. *Counseling Outcome Research and Evaluation, 4*(2), 75-98. https://doi.org/10.1177/2150137813494766

Oluwatayo, J. A. (2012). Validity and reliability issues in educational research. *Journal of Educational and Social Research, 2*(2), 391-400. https://doi.org/10.5901/jesr.2012.v2n2.391

Opara, I. M., & Magnus-Arewa, E. (2017). Development and validation of mathematics achievement test for primary school pupils. *British Journal of Education, 5*(7), 47-57. http://www.eajournals.org/wp-content/uploads/Development-and-Validation-of-Mathematics-Achievement-Test-for-Primary-School-Pupils.pdf

Parsian, N., & Dunning, T. (2009). Developing and validating a questionnaire to measure spirituality: A psychometric process. *Global Journal of Health Science, 1*(1), 2. https://doi.org/10.5539/gjhs.v1n1p2

Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? Critique and recommendations. *Research in Nursing & Health, 29*(5), 489-497. https://doi.org/10.1002/nur.20147

Recber, S., Isiksal, M., & Koc, Y. (2018). Investigating self-efficacy, anxiety, attitudes and mathematics achievement regarding gender and school type. *Anales de Psicología* [*Annals of Psychology*], *34*(1), 41-51. https://doi.org/10.6018/analesps.34.1.229571

Said, H., Badru, B. B., & Shahid, M. (2011). Confirmatory factor analysis (CFA) for testing validity and reliability instrument in the study of education. *Australian Journal of Basic and Applied Sciences, 5*(12), 1098-1103. https://www.researchgate.net/publication/261830668_Confirmatory_Factor_Analysis_Cfa_for_Testing_Validity_And_Reliability_Instrument_in_the_Study_of_Education

Salkind, N. J. (2010). *Encyclopedia of research design*. SAGE. https://doi.org/10.4135/9781412961288

Skaalvik, E. M., Federici, R. A., & Klassen, R. M. (2015a). Mathematics achievement and self-efficacy: Relations with motivation for mathematics. *International Journal of Educational Research, 72*, 129-136. https://doi.org/10.1016/j.ijer.2015.06.008

Sowtali, S. N., Yusoff, D. M., Harith, S., & Mohamed, M. (2016). Translation and validation of the Malay version of the stroke knowledge test. *Journal of Arrhythmia, 32*(2), 112-118. https://doi.org/10.1016/j.joa.2015.10.003

Suhr, D. D. (2006). *Exploratory or confirmatory factor analysis?* http://www2.sas.com/proceedings/sugi31/200-31.pdf

Sun, C. W., & PJU, K. D. (2009). Questionnaire translation and psychometric properties evaluation. *SEGi University College, 2*(2), 45-51. https://www.onlinereview.segi.edu.my/chapters/vol2_chap5.pdf

Taber, K. S. (2018). The use of Cronbach's alpha when developing and reporting research instruments in science education. *Research in Science Education, 48*(6), 1273-1296. https://doi.org/10.1007/s11165-016-9602-2

Taherdoost, H. (2016). Validity and reliability of the research instrument: How to test the validation of a questionnaire/survey in a research. *International Journal of Academic Research in Management, 5*(3), 28-36. https://doi.org/10.2139/ssrn.3205040

Tavakol, M., & Wetzel, A. (2020). Factor analysis: A means for theory and instrument development in support of construct validity. *International Journal of Medical Education, 11*, 245-247. https://doi.org/10.5116/ijme.5f96.0f4a

Tschannen-Moran, M., & Hoy, A. W. (2007). The differential antecedents of self-efficacy beliefs of novice and experienced teachers. *Teaching and teacher Education, 23*(6), 944-956. https://doi.org/10.1016/j.tate.2006.05.003

Usher, E. L., & Pajares, F. (2008). Sources of self-efficacy in school: Critical review of the literature and future directions. *Review of Educational Research, 78*(4), 751-796. https://doi.org/10.3102/0034654308321456

Usher, E. L., & Pajares, F. (2009). Sources of self-efficacy in mathematics: A validation study. *Contemporary Educational Psychology, 34*(1), 89-101. https://doi.org/10.1016/j.cedpsych.2008.09.002

Van Dinther, M., Dochy, F., & Segers, M. (2011). Factors affecting students' self-efficacy in higher education. *Educational Research Review, 6*(2), 95-108. https://doi.org/10.1016/j.edurev.2010.10.003

Volkmar, F. R. (2013). *Encyclopedia of autism spectrum disorders*. Springer. https://doi.org/10.1007/978-1-4419-1698-3

Yurt, E. (2014). The predictive power of self-efficacy sources for mathematics achievement. *Eğitim ve Bilim* [*Education & Science*], *39*, 176. https://doi.org/10.15390/EB.2014.3443

Zamanzadeh, V., Ghahramanian, A., Rassouli, M., Abbaszadeh, A., Alavi-Majd, H., & Nikanfar, A.-R. (2015). Design and implementation content validity study: Development of an instrument for measuring patient-centered communication. *Journal of Caring Sciences, 4*(2), 165. https://doi.org/10.15171/jcs.2015.017

Zhang, B., Li, Y. M., Li, J., Li, Y., & Zhang, H. (2016). The revision and validation of the academic motivation scale in China. *Journal of Psychoeducational Assessment, 34*(1), 15-27. https://doi.org/10.1177/0734282915575909

Zhou, M., & Brown, D. (2015). *Educational learning theories* (2nd Edn.). Education Open Textbooks. https://oer.galileo.usg.edu/education-textbooks/1