INTERNATIONAL JOURNAL OF ACADEMIC RESEARCH IN BUSINESS & SOCIAL SCIENCES
HRMARS MANAGEMENT ACADEMIC RESEARCH SOCIETY EXPLORING INTELLECTUAL CAPITAL
www.hrmars.com
ISSN: 2222-6990

A Confirmatory Factor Analysis on Teachers' Sense of Efficacy Scale for In-Service Teachers in Malaysia

Elia Md Johar

To Link this Article: http://dx.doi.org/10.6007/IJARBSS/v12-i9/14712 DOI:10.6007/IJARBSS/v12-i9/14712

Received: *09* July 2022, ***Revised:*** *12* August 2022, ***Accepted:*** *26* August 2022

Published Online: *07* September 2022

In-Text Citation: (Johar, 2022)

To Cite this Article: Johar, E. M. (2022). A Confirmatory Factor Analysis on Teachers' Sense of Efficacy Scale for In-Service Teachers in Malaysia. *International Journal of Academic Research in Business and Social Sciences*, *12*(9), 1028 – 1045.

Vol. 12, No. 9, 2022, Pg. 1028 – 1045

http://hrmars.com/index.php/pages/detail/IJARBSS | **JOURNAL HOMEPAGE**

Full Terms & Conditions of access and use can be found at
http://hrmars.com/index.php/pages/detail/publication-ethics

INTERNATIONAL JOURNAL OF ACADEMIC RESEARCH IN BUSINESS & SOCIAL SCIENCES
HRMARS MANAGEMENT ACADEMIC RESEARCH SOCIETY
EXPLORING INTELLECTUAL CAPITAL
www.hrmars.com
ISSN: 2222-6990

A Confirmatory Factor Analysis on Teachers' Sense of Efficacy Scale for In-Service Teachers in Malaysia

Elia Md Johar
Academy of Language Studies, Universiti Teknologi MARA(UiTM), Malaysia
Email: elia@uitm.edu.my

Abstract
This study presents the construct-related evidence of the Teachers' Sense of Efficacy Scale (TSES) developed by Tschannen-Moran and Woolfolk Hoy (2001) for in-service teachers in the Malaysian context. The under-researched validation of the TSES in the local context warranted an investigation into unrevealing the scale's psychometric properties. Thus, this cross-sectional study sought to determine the factor structure of the scale where factor loadings of each efficacy source contribution were determined. The 24-item scale, which drew on its sources from student engagement, instructional strategies, and, classroom management was explored through a confirmatory factor analysis method. The analysis demonstrated that a 9-item scale of the three-factor structure model well fitted the data of the local setting. The study recommends improvement on the items so that they could better represent the underlying constructs to match the culture within a different educational setting. This implicates the use of the scale in measuring teacher efficacy for in-service teachers in Malaysia as it was empirically proven to have factor stability.
Keywords: Teachers' Sense of Efficacy Scale, Confirmatory Factor Analysis, Student Engagement, Instructional Strategies, Classroom Management

Introduction

Teacher efficacy remains instrumental in determining teacher as well as school effectiveness. Tschannen-Moran (1998) defines teachers' sense of efficacy as 'the teacher's beliefs in his or her capability to execute courses of action required to accomplish a specific teaching task in a particular context' (p.22). Research has shown that teachers with high efficacy would positively leave a significant impact on student's achievement as they are in a better position to meet students' needs (Henson et al., 2001; Kim & Seo, 2018; Witcher et al., 2002) and can enhance student learning and motivation (Lancaster, 2014). Teacher efficacy helps bring about desired outcomes of student engagement and learning even among those students who may be difficult or unmotivated (Tschannen-Moran et al., 1998) and accounts for individual teacher differences in effectiveness. Teachers who believe in effective teaching, and their teaching abilities, should persist longer, provide a greater academic focus in the classroom and exhibit different types of feedback than those who have lower expectations

(Gibson & Dembo, 1984), and teach in a way that demonstrates that beliefs (Penrose et al., 2007).

There is always a growing interest to verify whether in-training teachers or beginning teachers have developed classroom teaching skills to be effective with their students or whether experienced teachers have acquired effective teaching skills over years. Besides teaching certification or licensure tests and performance assessments, there are still inadequate measures of teachers' effectiveness in contributing to student achievement (Darling-Hammond, 2010). What aggravates more is in the 21st century, more sophisticated forms of teaching are needed to meet the demand of developing student competencies, such as mastery of challenging content, critical thinking, complex problem solving, and, effective communication and collaboration which warrant refining teachers' pedagogical methods (Darling-Hammond et al., 2017) and thus indicating a dire need for the assessment of teachers efficacy to shed some insights into their beliefs in meeting such accountability. This has largely driven many parties including academic scholars to start validating and using teachers' sense of efficacy scale (TSES) for measuring teachers' effectiveness as teacher efficacy plays an important role in teachers' effectiveness (Barni et al., 2019; Sehgal et al., 2017).

TSES has been one of the dominant instruments in measuring teacher efficacy and its related areas and numerous studies on Tschannen-Moran and Woolfolk Hoy's (2001) TSES has been carried out in different educational settings (Covvarubias & Lira, 2016; Ling et al., 2015; Salas-Rodriguez, 2021; Valls et al., 2020), however, there is a lack of research on the subject in the Malaysian context (Ling et al., 2015; Khairani & Makara, 2020). This thus has warranted more investigations on examining the scale's psychometric properties for various groups of teachers in the local educational setting.

Objectives of the Study

The study aimed to:

(a) To examine the factor structure of the Teachers' Sense of Efficacy Scale
(b) To determine the inter-relationships between efficacy for student engagement, instructional strategies, and classroom management.

Literature Review

Teacher Efficacy Construct and the Scale Development

Bandura (1997), in his socio-cognitive theory, conceptualizes 'self-efficacy' as a judgment of capability that can affect and direct an individual's thought processes, the level, and persistence of motivation, and affective states towards the engagement of a particular behaviour. In other words, it is the belief of one's capabilities or mastery experience to organize and execute the courses of action required to produce given attainments. Dale and DiBenedetto (2021) describe it further as a personal, motivational construct that can affect and be influenced by behaviors and social or environmental variables, including choices, effort, persistence, and achievement of an individual.

Between 1970 and 2000, numerous researchers went through extensive reviews of literature before arriving at the conception of teacher efficacy. In essence, teacher efficacy was coined either on Bandura's or Rotter's social learning theory or both. According to Dembo and Gibson (1985), the teacher efficacy construct was first conceptualized by Berman and McLaughlin (1977) who did two RAND Corporation evaluation studies on teacher efficacy and discovered that teachers' sense of efficacy was positively related to the percentage of project goals achieved, amount of teacher change, the continuation of both project methods and

materials and, improved student performance. The conceptualization was grounded in Rotter's social learning theory (Kleinsasser, 2014; Skaalvik & Skaalvik, 2007; Tschannen-Moran & Hoy, 2001) which postulates one's self influences behavior which is directed by his or her internal locus of control (Rotter, 1966). Gibson and Demb (1984) conducted a construct-validation study on teacher efficacy where they developed a 30-item Teacher Efficacy Scale and found two substantial factors that corresponded to Bandura's 2-factor model of self-efficacy theory. More than a decade later, Tschannen-Moran et al. (1998) developed a two-dimensional model of teachers' sense of efficacy with a focus on teacher performance in the classroom context, incorporating teaching tasks and their context, and teachers' self-perception of teaching competencies. These are among the prominent efforts made in the conceptualization and operationalization of teacher efficacy.

Tschannen-Moran et al (1998) introduced an eclectic measure of teacher efficacy that was theoretically based on both competence and contingency which are both agent-means and means-ends relations. They incorporated Bandura's four sources of efficacy—cognitive processing, analysis of teaching task, assessment of personal teaching competence, and the consequences of teacher sense of efficacy (TSE) in their revised model of TSES**.** The model was further improved in 2001 by specifying the scope of teacher's competencies by applying personal competence and analysis of tasks in terms of resources and constraints in particular teaching contexts and teacher's assessment of their competence across the wide range of tasks and activities they are asked to perform (Friedman & Kass, 2002). This led to the formulation of a 3-dimensional instrument that contained the following efficacy dimensions – student engagement, instructional strategies, and classroom management.

The first dimension of the teacher efficacy construct in TSES is student engagement. Efficacy for student engagement refers to teachers' beliefs about their abilities to bring about desired outcomes of student engagement and learning (Bandura, 2006). It comprises both behavioural and emotional components. Students who can engage in learning show sustained behavioral involvement in learning activities and positive emotions.

Efficacy for instructional strategies refers to teachers' conceptions in their instructional practices on assessments, teaching, learning, and curriculum to promote students' thinking. It gauges the strength of teachers' beliefs regarding their ability to implement alternative teaching strategies and to use a variety of assessment strategies in the classroom (Tschannen-Moran et al., 1998). In addition, it gauges teachers' level of confidence in responding to difficult questions posed by the students and providing an appropriate challenge to more capable students

Efficacy for classroom management encompasses strategies aimed at increasing or, encouraging desirable student responses through praise, encouragement, attention, and rewards. Previous research reveals that this dimension is crucial before learning as failure to deter arising classroom problems associated with a misdemeanor can affect the teaching and learning process (Taxer et al., 2018; Valente et al., 2019).

The Long and Short Forms of TSES

The original Teachers' Sense of Efficacy (TSES) scale was developed in two formats: a 24-item long form and a 12-item short form (Tschannen-Moran et al., 1998; Tschannen-Moran & Hoy, 2001) which were designed to measure the degree to which a teacher perceives they can affect student performance. The former has 8 items per sub-scale while the former, the number of items per sub-scale was reduced to 4 items each. Tschannen-Moran and Woolfolk Hoy (2001) used exploratory factor analysis to validate the scale.

Validation of both scales identified three factors of teacher efficacy in student engagement, instructional strategies, and classroom management and thus concluded that they had a unified and stable factor structure (Tschannen-Moran & Hoy, 2001; Valls & Bonvin, 2015). Tschannen-Moran and Hoy (2001) reported that both scales could be reasonably valid and reliable as the psychometric properties of the short form of the TSES were nearly identical to those of the long form and recommended the short form was for pre-service teachers and the long one for in-service teachers when running a factorial analysis on both scales. Valls and Bonvin (2015) in their review of the scales argue that the short form should be preferred over the long version on the ground that the latter had redundant items and often showed high reliability. He also reported that the factor structure of the long version has been empirically proven in studies in North America and Europe (e.g. Klassen & Chiu, 2010; Swanson, 2012).

Validation of Teachers' Sense of Efficacy Measures

Numerous studies had attempted to validate the TSES scale, either using a long or a short version or both with factorial analysis approaches for in-service and pre-service teachers of different countries and languages (Salas-Rodriguez et al., 2021). Monteiro and Forlin (2020) in their validation study on the 12-item Chinese version scale as a measure of teacher self-efficacy for inclusive education in Macao used a factorial analysis with 118 pre-service teachers. Exploratory and confirmatory analysis indicated that the two-factor structure was not supported, however, after re-specification, a stronger 9-item of the two-factor scale was identified. In France, Valls et al (2020), with the same aim, used a short form of the French-translated TSES with a sample of 283 primary and secondary teachers. He reported that the scale had the same factor structure identical to the original version and demonstrated reasonable psychometric properties. Although both studies used the short form of TSES which had been translated into different languages with different types of teachers, the scale was remarkably proven to support the original factor structure.

Regarding the concern of the study which was the long form of TSES, many researchers found the factor structure of the 24-item scale was reliable and valid and recommended it as a measure of teacher efficacy. In Chile, using a sample of 544 in-service teachers, Covvarubias and Lira (2016) adapted the scale by reducing the scale from 24 to 17 items and grouped them by four different factors: the three original factors and a new one named *efficacy in attending to the uniqueness of students.* CFA results indicated that the adapted scale was valid, reliable, and explanatory of the self-efficacy in their context. In Mexico*,* Salas-Rodriguez et al. (2020) studied the long TSES psychometric properties with 190 primary and secondary Mexican teachers from 25 private schools. The confirmatory factor analysis (CFA) results indicated that the Spanish TSES had a three-factor correlated structure, identical to that of Tschannen-Moran and Woolfolk Hoy (2001) for in-service teachers. However, the RMSEA value (0.102) was higher than expected, and the TLI value (0.886) was just below the critical threshold (0.08 and 0.9, respectively). In Switzerland, with in-service teachers, Valls et al (2020) demonstrated the same CFA results where the three-factor structure was maintained but few items were dropped from the scale. Evidently, in the above studies, the long scale was reliable and valid.

Similarly, in the local setting, there were two studies so far that dealt with determining the psychometric properties of the long one. Ling et al (2015) used confirmatory factor analysis (CFA) on the long TSES to assess the uni-dimensionality, validity, reliability, and fitness of the teacher efficacy model with 349 in-service Malaysian secondary school teachers. The fully Bahasa Melayu translated 24-item scale was reduced to 13 items; 4 items – Student engagement factor, 4 instructional strategies items and, 5 items in Classroom Management

dimension were retained. The study concluded that the scale was best explained by the three-factor structure with 13 items, showing optimal convergent validity and discriminant validity. Khairani and Makara (2020) had 191 in-service and 122 pre-service Malaysian teachers sampled together replicating what had been done by (Tschannen-Moran and Hoy, 2001). They transformed the scale into a bilingual 1-5 Likert scale and tested the scale with the one-factor model, and the three-factor model, and, found that the three-factor model was a better fit. Two items (Q1 and Q2) did not satisfactorily load on the Efficacy for Student Engagement factor but Q5 best fit the factor. Hence, they concluded that the scale could be reliably used for other related studies.

Based on the above discussion, the study formulated the following hypotheses to be tested:

H01: The hypothesized structure fits the data well.

H02: Efficacy for student engagement, instructional strategies, and classroom management are not correlated.

The proposed three-common factor structure model for the teacher sense of efficacy construct is shown in Figure 1. Items or indicators from each scale were assumed to load only on their latent variables. The model also demonstrated covariance among the three efficacy judgments, namely, Student Engagement, Instructional Strategies, and Classroom Management.

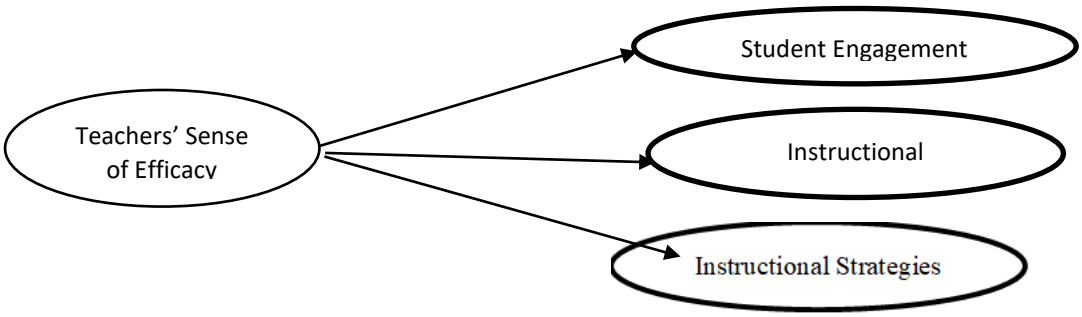

Figure 1 The proposed three-factor structure of Teachers' Sense of Efficacy

Method

Research Design

The study was cross-sectional survey research involving administering the long TSES scale, originally developed by Tschannen-Moran et al. in 1998 and revised by Tschannen-Moran and Woolfolk Hoy in 2001. It was quantitative as it concerned construct validation of the instrument through determining the factor structure of the TSES scale.

Sample

Survey questionnaires were distributed to 125 conveniently selected respondents ranging from primary and secondary teachers in Johor, Perak, and Selangor due to the review of the literature made by Salas-Rodriguez (2021) who identified the similar sampling technique employed by Tschannen-Moran et al (1998) and other researchers. A total of 160 questionnaires were sent out for collecting data, however, the response rate was only 78%. For a confirmatory factorial analytic approach to be performed, Anderson and Gerbing (1988) recommend a minimum of 150 while Chou and Bentler (1995) suggest that a reasonable but practical sample is 200. Although the inadequacy of the sample size would affect the results,

the study was still carried out as several studies albeit having enough samples, similar results were obtained with few inaccuracies (Monteiro & Forin, 2020; Salas-Rodriguez, 2021)

The composition of the sampled in-service teachers was as follows: 36 teachers at a full boarding school, 70 were from two regular secondary schools and 19 were primary school teachers. The sampling comprised 27 males (21.6%) and 98 females (78.4%) with an average age of approximately 33 years. Most of them (85%) had tertiary education.

Instrument

The 24-item Teachers' Sense of Efficacy Scale (TSES), developed by Tschannen-Moran and Woolfolk Hoy (2001) was used for the validation study. The Cronbach Alpha values for the sub-scales were high at .87, .90, and .91 respectively (Tschannen-Moran & Woolfolk Hoy, 2001).

Although the scale was not adapted, two minor changes were made to the scale to fit the local context, particularly the sampled teachers. First, a 9-point Likert scale was modified to a 5-point continuum, ranging from 1 – Nothing, 2 – Very Little, 3 – Some Influence, 4 – Quite A Bit, and 5 – A Great Deal. The reason for the change was to follow the common practices of the item construction involving the Likert scale rating among Malaysian researchers and for convenient use both in data entry and analysis. Lastly, the English version scale was transformed into a bilingual version – each item of the sub-scales was followed by the Bahasa Melayu translated version of the item. The translated items were checked by a colleague who taught Bahasa Melayu programme. Rephrasing the items was not necessary as this could deter the achievement of the research objectives of the study. Having a bilingual TSES scale in the local setting would help enhance the understanding of each item and thus, resulting in more accurate responses and data.

The dimensions of the scale in the study were labelled as follows: Efficacy in Student Engagement - Factor 1, Efficacy in Instructional Strategies - Factor 2, and Efficacy in Classroom Management -Factor 3. For easy reporting, the three scales were labelled efficacy factors, and each item of the sub-scales or, factors represented an indicator. It was assumed that for each indicator under the same factor, they would measure the same construct. Hence, the eight indicators for each factor were assumed to have the same respective construct.

Data Collection

As the study was preliminary and conducted on a small scale, the school administrators of one primary school and two government secondary schools were contacted privately over the phone and sought their discretion to allow for online data collection. Similarly, the same move was made for data collection at a full boarding secondary government school assisted by a school administrator. Of 160 questionnaires, 125 were returned at a response rate of 75 percent. The informed consent from each teacher was gained before the scale administration.

Data Analysis

A confirmatory factorial analytic approach employed in the study incorporated a model development approach followed by model trimming. It began with a hypothesized model and moved on by dropping a path one by one to find the most parsimonious model by relying on descriptive fit indices. To arrive at a well-fitting model, the conclusion, the study would employ the AMOS, model fitting program.

Data were analyzed in three steps. First, checking the statistical assumptions for normality was done by running descriptive statistics to get skewness and kurtosis. Second, a

confirmatory factor analysis with maximum likelihood estimation was run to test the proposed factorial structure of the TSES. Goodness fit tests were entered as well to determine the model fit. A review of factor loadings on indicators would lead to the elimination of poorly loaded items. Third, exploratory factor analysis was entered to determine whether the fit of the best fitting model in Step 1 could be improved. For this purpose, the standardized regression coefficients were first examined to look for those which did meet the criterion. Next, the modification indices were examined to search for items that were loaded on two factors. The model was then respecified by dropping paths with poor loadings. The process would go on by running the goodness-fit tests till the well-fitting model was found.

The evaluation of the model fit I would be based on the Chi-Square Likelihood Ratio, the Root Mean Square Error of Approximation (RMSEA), the Tucker-Lewis Index (TLI), and the Comparative Fit Index (CFI). The Chi-Square Likelihood Ratio should show its insignificance at $p > .05$ or more stringent one, $p > .01$. A RMSEA value of .08 or less indicates the model fit (Hair et al., 2010). Alternatively, CMIN/df with a value of between 2 and 5 would be considered acceptable and the possible values of CFI and TLI range from zero to 1, with values close to one, indicating a good fit (Arbuckle & Wothke, 1999).

Results

The normality of the data was tested and data is considered to be normal if skewness is between -2 to +2 and kurtosis is between -7 to +7 (Bryne, 2010; Hair et al., 2010). Both skewness and kurtosis results demonstrated a normal distribution of data. The kurtosis values for the indicators (items) ranged from -.86 to 1.75 and this indicated that the values were within the acceptable range for normality. Similarly, the skewness values ranging from -.79 to -.20 also fell within the normality range.

The internal consistency analysis of the bilingual 24-item TSES showed very strong reliability at .91 from 125 responses. The three factors also had high-reliability coefficients with Cronbach's alpha values of efficacy in student engagement .81, efficacy in classroom management .85, and, efficacy in instructional strategies .83. Thus, the scale still had good reliability values although the sample size of the study was inadequate for factor analysis.

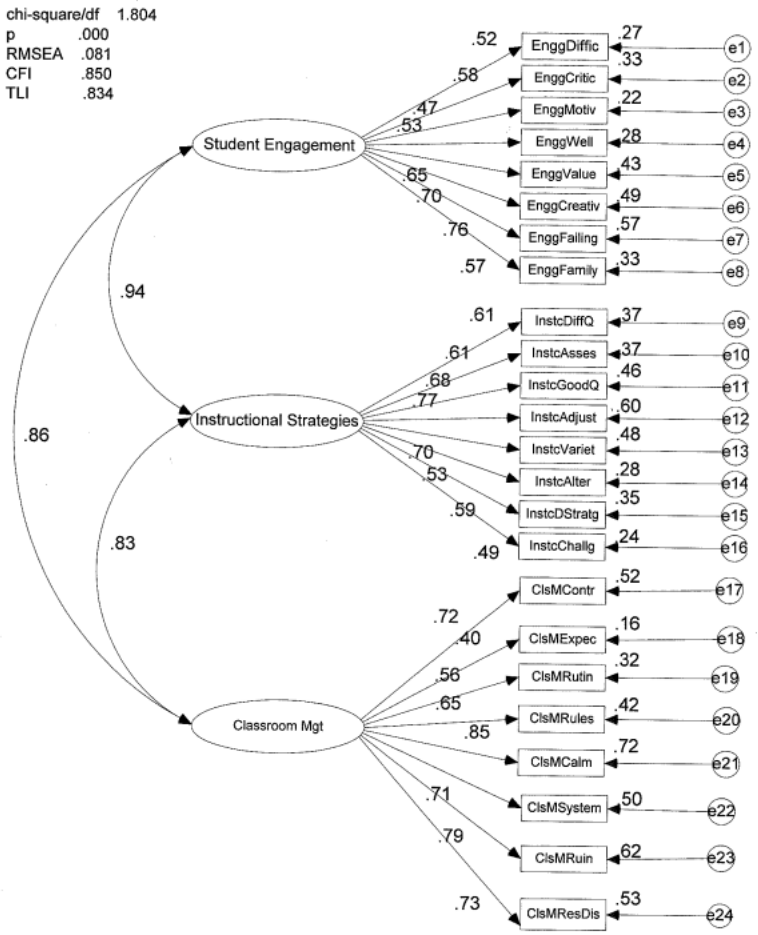

Figure 2 The hypothesized model of the Teachers' Sense of Efficacy Scale

The results of the overall fit of the model were not encouraging as displayed in Table 1. Inspection of the goodness fit test which rendered the fit of the model as x^2 (249, N=125) = 449.1225, $p < .05$ indicated that the Chi-square goodness of fit was significant. This indicated that the model's covariance structure was significantly different from the observed covariance matrix. Therefore, H01 was rejected because the hypothesized model lacked fit or was not consistent with the data. In addition, an RMSEA value of .081 signified the lack of fit which was due to the significant discrepancy in the population covariance matrix. This was further substantiated by both CFI and TLI values which were .85 and .83 respectively, below the threshold of .90. Based on the recommendations by Bentler and Bonett (1980), the fit of a model was considered to be acceptable if TLI and CFI exceeded .90. All these indices indicated a lack fit of a measurement model and thus, required a model re-specification.

Table 1
Fit Indices Confirmatory Factor Analysis Results

Fit index	**Statistics**
Chi-square	449.123
RMSEA	.081
CFI	.850
TLI	.834

Revised Model 1

To improve the model fit, an inspection of the pattern of standardized loadings in Table 2 indicated one possible reason for the model's lack of fit. The squared multiple correlations (SMC) for indicator *EnggDiffi*c, which significantly loaded on Factor 1 was .27. This indicated that the factor extracted only about 27% of the variance in the indicator, thereby affecting the reliability of the indicator. Similarly, under the same factor, other SMCs that posed a problem of model fit were those of *EnggCritic* (.33), *EnggMotiv* (.22), *EnggWell* (.28), and *EnggFamily* (.33). Other similar patterns also could be observed for the other two factors, involving the following indicators: *InstcDiffQ* (.37), *InstcAssess* (.37) *InstcAlter* (.28), *InstcDStratg* (.35), *InstcChallg* (.24), *ClsMExpect* (.16), *ClsMRutin* (.32), and *ClsMRules* (.42). A common rule of thumb is that the indicators or items should have a loading of .7 to judge the set reliable. However, there were indicators like *EnggValue* (.43) and *InstcGoodQ* (.46) having SMC values below .50 retained as the minimum number of indicators that should be allowed to stay is 3 indicators. Therefore, Factor 1 and Factor 2 would have three indicators to measure the construct and five indicators remained under Factor 3.

Table 2
Standardized Loadings and Squared Multiple Correlations: Hypothesized and Revised Models

Indicators (Labels)	Model			
	Hypothesized		**Revised**	
	Loading	SMC	Loading	SMC
Factor 1				
How much can you do to get through the most difficult students? (*EnggDiffic)*	.52	.27	Deleted	
How much can you do to help your students think critically? (*EnggCritic*)	.58	.33	Deleted	
How much can you do to motivate students who show low interest in schoolwork? (*EnggMotiv*)	.47	.22	Deleted	
How much can you do to get students to believe they can do well in school work? (EnggWell)	.53	.28	Deleted	
How much can you do to help your students to value learning? (*EnggValue)*	.66	.43	.60	.36
How much can you do to foster student creativity? (EnggCreativ)	.70	.49	.69	.47
How much can you do to improve the understanding of a student who is failing? (*EnggFailing)*	.76	.57	.80	.64

How much can you assist families in helping their children do well in school? ***(****EngFamily)*	.57	.33	Deleted	
Factor 2				
How well can you respond to difficult questions from your students*? (InstcDiffQ)*	.61	.37	Deleted	
How much can you assess student comprehension of what you have taught? (*InstcAssess)*	.61	.37	Deleted	
To what extent can you develop good questions for your students? *(InstcGoodQ)*	.68	.46	.66	.44
How much can you do to adjust your lessons to accommodate individual differences among students? (*InstcAdjust)*	.77	.60	.80	.64
How much can you use a variety of assessment strategies? (*InstcVariet)*	.70	.48	.68	.47
To what extent can you provide an alternative explanation or example when students are confused? (*InstcAlter)*	.53	.28	Deleted	
How well can you implement different strategies in your classroom? (*InstcDStratg)*	.59	.35	Deleted	
How well can you provide appropriate challenges for very capable students? *(InstcChallg)*	.49	.24	Deleted	
Factor 3				
How much can you do to control disruptive behavior in the classroom? (*ClsMContr)*	.72	.52	.71	.51
To what extent can you make your expectations clear about student behavior? (*ClsMExpect)*	.40	.16	Deleted	

How well can you establish routines to keep classroom activities running smoothly? (*ClsMRutin)*	.56	.32	Deleted	
How much can you do to get students to follow classroom rules? (*ClsMRules)*	.85	.42	Deleted	
How much can you do to calm a student who is disruptive or noisy? (*ClsMCalm)*	.71	.72	.86	.73
How well can you establish a classroom activities management system with a student of different behaviours? (*ClsMSystem)*	.79	.50	.72	.53
How well can you keep a few problematic students from ruining an entire lesson? (*ClsMRuin)*	.68	.62	.79	.63
How well can you respond to disobedient students? (*ClsMResDis)*	.73	.53	.73	.53

Items to be retained for each factor are as follows:(1) 3 items to be retained under Factor 1 were *EnggValue* (How much can you do to help your students to value learning?), *EnggCreativ* (How much can you do to foster student creativity?) and, *EnggFailing* (How much can you do to improve the understanding of a student who is failing?). (2) 3 items to be retained for Factor 2 (Instructional Strategies) were: *InstcGoodQ* (To what extent can you develop good questions for your students?), *InstcAdjust (*How much can you do to adjust your lessons to accommodate individual differences among students?) and, *InstcVariet* (How much can you use a variety of assessment strategies?) (3) As for Factor 3 (Classroom Management), 5 items to be retained were: *ClsMContr* (How much can you do to control disturbing behavior in the classroom?), *ClsMCalm* (How much can you do to calm a student who is disruptive or noisy?), *ClsMSystem* (How well can you establish a classroom activities management system with students of different behaviour?), *ClsMRuin* (How well can you keep a few problematic students from ruining an entire lesson?) and, *ClsMResDis* (How well can you respond to disobedient students?). They were loaded with very high values with .71, .86, .72, .79 and .73 respectively. Altogether, there were 11 items to be included in the revised model.

The fit indices of the revised model in Table 3 showed that the goodness fit of the model was x^2 (41, N=125) = 86.61, p =.000 (p < .05), indicating the significant Chi-square goodness of fit result. Therefore, the result again showed that H01 was rejected because the revised model lacked fit or was not consistent with the data.

Table 3
Fit Indices Confirmatory Factor Analysis Results

Fit index	**Statistics**
Chi-square	86.61
RMSEA	.095
CFI	.934
TLI	.912

Although CFI and TLI demonstrated better values and exceeded .90 which were .934 and .912 respectively, an RMSEA value of .095 did not meet the fit criterion. Hair et al. (20106) ruled out that "…simpler models and smaller samples should be subject to more strict evaluation (p.753)." and evidence of good fit of a similar case to the revised model which was as follows "…based on a sample of 100 respondents and a four-construct model with 12 indicator variables, evidence of good fit would include an insignificant x^2value, a CFI at least .97 and an RMSEA of .08 or lower (p.753)." Thus, the model still warranted re-specification as the Chi-Square and the RMSEA value did not show the model fit. There were still modification indices that were above the threshold value of 4.0. To reduce the overall chi-square value, the possible path that could be considered of dropping to improve the fit is Item 21 (*ClsResDis*) whose modification value was 7.368.

Revised Models 2 and 3

The model fit was proven to be improved substantially when the number of items started leveling for Classroom Management Factor. For the revised Model 3, there were four items left for Classroom Management and the correlation coefficients among the latent factors were not much different. The fit indices particularly the chi-square value have decreased greatly ($p<.01$) as shown in Table 9 but it still indicated a value above the criterion .08. CFI and TLI improved substantially as well.

Finally, the last revised model in Figure 2 indicated the best fit model as the fit indices met the criterion of a well-fitting model. The insignificant Chi-square, x^2 (24, N=125) = 41.74, $p > .01$ indicated that the null hypothesis failed to be rejected, suggesting the model fitted the data. An RMSEA value below .08 for the model also showed an acceptable range

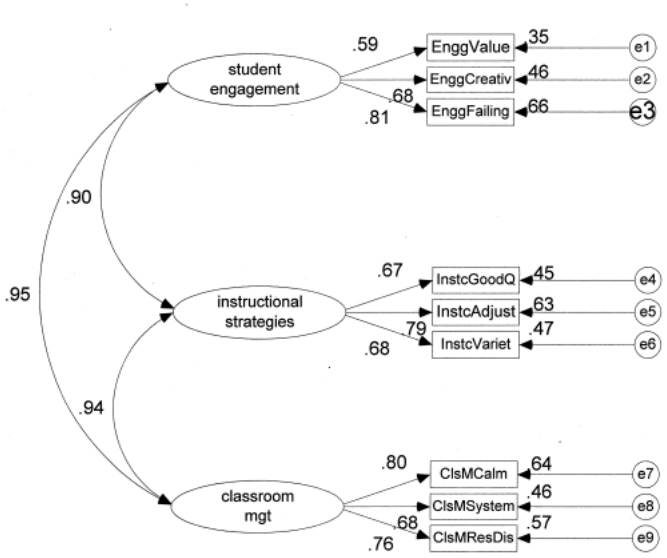

Figure 2: The last revised model for the Teachers' Sense of Efficacy Scale

for the model fit as well as CFI and TLI which both had exceeded .95. Since Bentler (1990) himself claimed the x^2distributed test statistics is problematic, Ullman (2001) suggested using relative Chi-square (Cmin/df) which only required a value of 2 to 5 to reflect good fit. In this context, in the revised Models 3 and 4, the relative Chi-Square values were 1.854 and 1.739 respectively which satisfactorily met the criterion of the model fit. Therefore, the 9-item Teachers' Sense of Efficacy model now did generate the data of the local setting.

Table 4
Summary of the Fit Indices for all the Models

N= 125	x^2	df	p	RMSEA	CFI	TLI
24-item model	449.123	249	.000	.081	.850	.834
11-item model	86.61	41	.000	.095	.934	.912
10-item model	59.3	32	.002	.83	.954	.936
9-item model	41.74	24	.014	.077	.964	.947

The results for inter-factor correlation are shown in Table 4 which indicated that the three efficacy factors were highly correlated with each other: Factor 1 with Factor 2 (.936), Factor 2 with Factor 3(.833), and Factor 1 with Factor 3 (.861). These high inter-factor correlations substantiated the expectation that the three factors were distinct and yet positively connected.

Table 5
Inter-factor Correlations

			Estimate
Student Engagement	⟺	Instructional Strategies	.94
Instructional Strategies	⟺	Classroom Management	.83
Student Engagement	⟺	Classroom Management	.86

Discussions

The study aimed to construct-validate the 24-item TSES, which established the psychometric properties of the scale within the local setting. The study confirmed that the three-factor structure measurement model was consistent with the data collected on the sampled in-service teachers in Malaysia. Although the analyses went through three revised models, the well-fitting model was eventually achieved with the total number of items being greatly reduced to a-9 item scale, indicating both versions of TSES were appropriate for both pre-service teachers and in-service.

Elimination of poorly fitted indicators and employing the modification indices as a guide to find the most parsimonious model did produce an adequate fit to the data. Model 2 was re-specified with 11 items, Model 3 with 10 items, and lastly, Model 4 was again equally categorized with 3 items for every subscale, reflecting the original version with 8 items for every subscale. The scale remained in a three-factor structure but in a short form. This finding concurred with the local validation studies by (Ling et al., 2015; Khairani and Makara, 2020). The two studies also demonstrated the factor stability of the long scale. Like in Khairani and Makara's study, the study also found 3 items with low loadings, two from efficacy for student engagement – 'How much can you do to get through to the most difficult students?', 'How much can you do to help your students think critically?' and one item from efficacy for class management 'To what extent can you make your expectations clear about student

behaviour?' and therefore, were removed. A comparison of the item deletion with Ling et al. (2015) could not be made as the items were not indicated properly in the labeling. Finally, the study also found strong positive correlations between factors, especially between efficacy for student engagement and efficacy for instructional strategies similar to Khairani and Makara's finding on inter-correlations between the efficacy factors.

It is noteworthy that many items under classroom management had high factor loadings, indicating that they were statistically significant indicators compared to other indicators under the other factors. This would call for revising items under the two efficacy factors.

The study also supports the assertion that teacher sense of efficacy is a multidimensional construct, the dimension being student engagement, instructional strategies, and classroom management. This noteworthy finding is in line with that of the original version of TSES and other previous studies and, thus, suggests that the measure has a consistent unified and stable factor structure (Khairani & Makara, 2020; Koehler, 2006; Ling et al., 2015; Monteirro & Forlin, 2020; Salas-Rodriquez, 2021).

Conclusions and Implications

The study concluded that the long TSES which contained 24 items had a stable factor structure and could be reliably used for other comparable studies in measuring teachers' effectiveness and specifically teachers' sense of efficacy. However, revising the removed items in a more related context of today's classroom after the Covid-19 pandemic was recommended.

The findings of the study have shed some implications for teacher efficacy beliefs in a Malaysian setting. The content of significant indicators represents what in-service teachers highly believe in particular in areas of student engagement, instructional practices, and classroom management. This, therefore, provides some information on how to enhance teacher effectiveness and detect the extent of teacher efficacy through task-specific areas.

The study has significantly contributed to the expansion of the literature on TSES validation and proven its relevant use in the Malaysian context, especially in assessing teacher efficacy for Malaysian in-service teachers. Furthermore, the findings have indicated through the scale's stable factor structure, that the efficacy sources proposed by Bandura (1997) are still applicable and appropriate to many kinds of educational settings, thus the measure of teacher efficacy could be pursued using the same scale even in different countries.

There may be several reasons why the hypothesized model could not easily be confirmed by its factorial structure. First, the sample of the study was heterogeneous. Second, the sample size itself was small, lesser than 200 which was an adequate sample size for a confirmatory factorial analytic approach to be performed. Third, the internal consistencies of the items themselves were already low. Finally, the content of certain items could not measure the same construct which might explain the low loadings. Thus, the minimum number recommended for CFA must be adhered to.

Since validation studies are abundant, the study calls for qualitative studies that emphasize a deeper focus on teachers' efficacy beliefs and sources, and for more longitudinal studies of teacher efficacy to provide more data on the development and stability of the teacher efficacy construct. These will help enhance the understanding of teacher efficacy of pre-service and in-service teachers.

References

Anderson, J., & Gerbing, D. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin*, 103, 411-423. http://dx.doi.org/10.1037/0033-2909.103.3.411

Arbuckle, J., & Wothke, W. (1999). *AMOS 4 user's reference guide*. Smallwaters Corporation.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. W. H. Freeman and Company.

Bandura, A. (2006). Toward a psychology of human agency. *Perspectives on Psychological Science, 1*(2), 164–180. https://doi.org/10.1111/j.17456916.2006.00011.x

Barni, D., Danioni, F., & Benevene, P. (2019). Teachers' self-efficacy: The role of personal values and motivations for teaching. *Front Psychol*. 31(4), 505-517. https://doi.org/ 10.3389/fpsyg.2019.01645.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, 107, 238-246.

Bentler, P. M., and Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, 88(3), 588.

Berman, P., & McLaughlin, M. W. (1977). *Federal Programs Supporting Educational Change, Volume VII: Factors Affecting Implementation and Continuation*. The RAND Corporation.

Byrne, B. M. (2010). *Structural equation modeling with AMOS: Basic concepts, applications, and programming*. Routledge.

Chou, C. P., & Bentler, P. M. (1995). Estimates and tests in structural equation modeling. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 37–55). Sage Publications, Inc.

Covarrubias, A., Apablaza, C. G., & Lira, M. C. (2016). Adaptación y validación del cuestionario sentimiento de autoeficacia en una muestra de profesores chilenos. *Univ. Psychol*. 15, 97–108. https://doi.org/10.11144/Javeriana.upsy15-2.avcs.

Dale, H. S., & DiBenedetto, M. K. (2021). Chapter four - Self-efficacy and human motivation. In Elliot, A.J (Eds.), *Advances in Motivation Sciences*. (pp. 153-179), ScienceDirect. https://doi.org/10.1016/bs.adms.2020.10.001.

Darling-Hammond, L. (2010). Evaluating t*eacher effectiveness: How teacher performance assessments can measure and improve teaching*. Centre of American Progress.

Darling-Hammond, L., Hyler, M. E., Gardner, M. (2017). *Effective teacher professional development.* Learning Policy Institute.

Dembo, M. H., & Gibson, S. (1985). Teachers' sense of efficacy: An important factor in school improvement. *The Elementary School Journal*, *86*(2), 173-184. https://doi.org/10.1086/461441.

Friedman, Y., & Kass, E. (2002) Teacher *self-efficacy: Concept and measurement*. Henrietta Szold Institute Press. (In Hebrew)

Gibson, S., & Dembo, M. (1984). Teacher efficacy: A construct validation. *Journal of Education Psychology*, 76, 569-582.

Hair, J., Black, W. C., Babin, B. J., & Anderson, R. E. (2010) *Multivariate data analysis* (7th ed.). Pearson Educational International.

Henson, R. K. (2001). *Teacher Self-Efficacy: Substantive implications and measurement dilemmas*. The invited keynote address was given at the annual meeting of the Educational Research Exchange, Texas A&M University, College Station, Texas

Khairani, A. Z., and Makara, K. A. (2020) Examining the factor structure of the teachers' sense of efficacy scale with Malaysian samples of in-service and pre-service teachers. *Pertanika Journal of Social Science and Humanities, 28*(1), pp. 309-323.

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

Kleinsasser, R. (2014). Teacher efficacy in teaching and teacher education. *Teaching and Teacher Education*, *44*, 168-179. https://doi.org/10.1016/j.tate.2014.07.007

Kim, K. R., & Seo, E. H. (2018). The relationship between teacher efficacy and student's academic achievement: A meta-analysis. *Social Behavior and Personality*, *46*(4), 529–540.

Koehler, J. R. (2006). *The measurement of teacher efficacy*. Poster presented at the 2006 Harvard Graduate School of Education Student Research Conference and International Forum. Cambridge, MA. www.icteams.umd.edu/HarvardHandout.pdf

Lancaster, J. (2014). School and classroom indicators of inclusive education. Measuring inclusive education. *International Perspectives on Inclusive Education*, 3, 227-245. https://doi.org/10.1108/S1479-363620140000003027

Ling, T. P., Akmaliah, Z. L. P., Asimirin, S., & Fooi, F. S. (2015). The validity and reliability of teacher efficacy: Revisited in Malaysia secondary schools. *Journal of Studies in Education, 5*(1), 1-34. https://doi.org./*10.5296/jse.v5i1.6802*

Monteiro, E., & Forlin, C. (2020). Validating the use of the 24-item long version and the 12-item short version of the Teachers' Sense of Efficacy Scale (TSES) for measuring teachers' self-efficacy in Macao (SAR) for inclusive education [version 1; peer review: awaiting peer review]. *Emerald Open Res* 2020, 2:36 https://doi.org/10.35241/emeraldopenres.13541.1

Penrose, A., Perry, C., & Ball, I. (2007). Emotional intelligence and teacher self-efficacy: The contribution of teacher status and length of experience. *Issues in Educational Research*, *17.* http://iier.org.au/iier17/penrose.html.

Rotter, J. B. (1966). Generalized expectancies of internal versus external control of reinforcements. *Psychological Monographs*, *80*(1), 1–28.

Salas-Rodriguez, F., Lara, S., & Martinez, M. (2021). Spanish version of the Teachers' Sense of Efficacy Scale: An adaptation and validation study. *Frontiers in Psychology*. https://doi.org/10.3389/fpsyg.2021.714145.

Sehgal, P., Nambudiri, R., and Mishra, S. K. (2017), Teacher effectiveness through self-efficacy, collaboration and principal leadership. *International Journal of Educational Management*, 31(4), 505-517. https://doi.org/10.1108/IJEM-05-2016-0090

Skaalvik, E. M., & Skaalvik, S. (2007). Dimensions of teacher self-efficacy and relations with strain factors, perceived collective teacher efficacy, and teacher burnout. *Journal of Educational Psychology*, 99(3), 611 - 625.

Swanson, P. (2012). Second/Foreign language teacher efficacy and its relationship to professional attrition, 68(1), 78-101, *The Canadian Modern Language Review*. https://doi.org/10.3138/cmlr.68.1.078

Taxer, J. L., Becker-Kurz, B., & Frenzel, A. C. (2018). Do quality teacher-student relationships protect teachers from emotional exhaustion? The mediating role of enjoyment and anger. *Social Psychology of Education, 22* (1), 209–226, https://doi.org/10.1007/s11218-018-9468-4

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

Tschannnen-Moran, M., & Woolfolk Hoy, A. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and Teacher Education*,*17*, 783-805.

Ullman, J. B. (2001). Structural equation modeling. In: B. G. Tabachnick, & L. S. Fidell (Eds.), *Using multivariate statistics*. Pearson Education.

Valente, S., Monteiro, A. P., & Lourenco, A. A. (2019). The relationship between teachers' emotional intelligence and classroom discipline management. *Psychology in the Schools*, 56 (5), 741-750, https://doi.org/10.1002/pits.22218.

Valls, M., & Bonvin, P. (2015). Auto-efficacitŽ des enseignants: quels outils dÕŽvaluation utiliser ? Mesure et ƒvaluation en ƒducation, 38-47. https://doi.org/10.7202/1036698a

Valls, M., Bonvin, P., & Benoit. (2020). Psychometric properties of the French version of the Teachers' Sense of Efficacy Scale (TSES-12f), *European Review of Applied Psychology*, 70(3), https://doi.org/10.1016/j.erap.2020.100551

Witcher, L. A., Onwuegbuzie, A. J., Collins, K. M. T., Witcher, A. E., James, T. L., & Minor, L. C. (2002). *Relationship between teacher efficacy and beliefs about education among preservice teachers*. Mid-South Educational Research Association. http://asstudents.unco.edu/students/AE-Extra/2006/10/Witcher.html