Validity Evidence for Teacher Self-efficacy (TSE) Scale

Tuba Gezer, Ph.D.
The Ohio State University, USA, https://orcid.org/0000-0001-9819-4146

Angela Watson, Ph.D.
Johns Hopkins University, USA, https://orcid.org/0000-0003-3695-1273

Alanna Bjorklund-Young, Ph.D.
Johns Hopkins University, USA, https://orcid.org/ 0000-0002-9744-8519

Al Passarella
Johns Hopkins University, USA, , https://orcid.org/0009-0005-2902-4297

Abstract: This study aimed to examine the validity and reliability evidence of Teacher Self-efficacy (TSE) scale (Tschannen-Moran & Hoy, 2001) for teacher in K-12 US schools. Particularly, our study gathered evidence for the internal structure validity, convergent validity, criterion validity and reliability of the TSES in US educational context. The survey was administered to Pre-K-12th grade teachers from a large school district in the Southwestern of the United States. The total sample size was 1,418. The majority of the sample was female (73.9%) and White (73.8%). Nearly half of the sample had BA in education (48.9%) and about 28% of them have master's degree in education. We conducted Confirmatory Factor Analysis (CFA) on Mplus. Data was screened, and all assumptions, normality, outliers and adequate sample size were met. We used CFA for internal structure, correlation for convergent validity evidence, and Analysis of Variance (ANOVA) for criterion validity evidence. Cronbach's Alpha and McDonald's Omega were used for the reliability. Our results showed that the data fits the model well based on the goodness-of-fit statistics. This study provides evidence of the validity and reliability of TSE using this sample.

Keywords: Teacher self-efficacy, Validity evidence, Reliability evidence

Citation: Gezer, T., Watson, A., Bjorklund-Young, A., & Passarella A. (2023). Validity Evidence for Teacher Self-efficacy (TSE) Scale. In M. Shelley, O. T. Ozturk, & M. L. Ciddi, *Proceedings of ICEMST 2023--International Conference on Education in Mathematics, Science and Technology* (pp. 212-225), Cappadocia, Turkiye. ISTES Organization.

Introduction

Teacher Self-efficacy (TSE) scale was developed by Tschannen-Moran and Hoy (2001) and has been widely

used in literature. According to Tschannen-Moran and Hoy's (2001), teacher efficacy refers to a teacher's assessment of their own ability to achieve the desired results in terms of student participation and learning. TSE scale was developed by Tschannen-Moran and Hoy (2001) and has been the most commonly used scale for teacher self-efficacy studies (Ma et al., 2019). Their research was consistent of three parts. In the first part, 52 items were developed and reduced to 32 items using principal-axis factoring with a sample of 224 teachers. For the second part, the scale was reduced to 18 items using the same analysis and another sample of 217 teachers. Lastly, they created a short (12 items) and long (18 items) of the TSE scale using a sample of 410 teachers. In addition to the validity evidence provided by Tschannen-Moran and Hoy (2001), Nie et al. (2012) also validated the construct in Singapore.

Specifically, they conducted CFA and examined the distinction between the sub-constructs, which are teacher efficacy in relation to: instructional strategies, motivation, and classroom management. Researchers reported high correlations between teacher self-efficacy and teaching strategies, indicating strong prediction validity. While there were some international studies (Fathi & Rostami 2018; Klassen et al. 2009; Nie et al., 2012) that examine the validity of teacher self-efficacy scale, there is a need to validate the scale in the US educational context. This study aimed to examine the validity and reliability evidence of TSES for teacher in K-12 US schools. Particularly, our study gathered evidence for the internal structure validity, convergent validity, criterion validity and reliability of the TSES in US educational context.

Factors Associated with Teacher Self-efficacy: Job Satisfaction, Teacher-Student Interaction and Teaching Experience

Job satisfaction can generally be defined as having a positive reaction to the workplace (Worrell et al., 2006). Within the field of education, research suggests that teacher job satisfaction may come from several different sources. For example, research suggests that positive social relationships are more likely to increase teacher job satisfaction (Sylvia & Hutchinson, 1985) and might also play a crucial role for teachers (Van Droogenbroeck et al., 2014). In addition, positive relationships with colleagues, parents, students are related to teacher satisfaction (Cano-Garcia et al., 2005; Gavish & Firedman, 2010; Skaalvik & Skaalvik, 2011). A meta-analysis by Kasalak and Dagyar (2020) used over 100 from 50 countries independent data to examine the relationship between job satisfaction and teacher self-efficacy. They concluded there is a positive correlation between teacher self-efficacy and job satisfaction.

Previous research indicated that there is a connection between teacher self-efficacy beliefs are linked with positive interactions with students in the classroom (Bloom & Peters, 2012; Summers et al. 2017; Siwatu & Starker, 2010). Davis et al. (2017) investigates the effects of teachers' efficacy beliefs on students' perceptions of the quality of their relationship with their teacher. The study finds that students are more likely to perceive a positive relationship with their teacher when the teacher has a strong sense of efficacy, indicating confidence in their ability to teach effectively.

Literature suggest that teacher self-efficacy increase as they gain teaching experiences (George et al., 2018; Wolters & Daugherty, 2007). Gale et al. (2021) examined teacher self-efficacy based on teaching experience and found a significant difference between first year teacher, novice teacher and career teachers where first year teacher had statistically significantly lower teacher self-efficacy.

Method

Data Collection and Participants

The survey was administered to Pre-K-12th grade teachers between March 14th and April 1st, 2022 from a large school district in the Southwestern of the United States. Of the 3,264 teachers who received the survey, 2,260 teachers responded, for a response rate of 69.2%. If a participant responded at least 50% of the survey, the responses were analyzed. Out of 2,260 surveys, 62.7% (n=1,418) met this completion criteria and are included in the final sample size in this analysis. Table 1 shows the demographic characteristics of the sample. The majority of the sample was female (73.9%) and White (73.8%). Nearly half of the sample had BA in education (48.9%) and about 28% of them have master's degree in education.

Table 1. Demographic Characteristics of the Participants

	N	%
Gender		
Female	1,048	73.9%
Male	281	19.8%
Other	20	1.4%
Missing	69	4.9%
Race/Ethnicity		
Native American	8	0.6%
Asian	26	1.8%
Black/African American	15	1.1%
Hispanic/Latino	146	10.3%
White	1,046	73.8%
Multi-racial	47	3.3%
Other	54	3.8%
Missing	76	5.4%
Education		
BA in Education	693	48.9%
University based post-BA program	168	11.8%

Master's in education	398	28.1%
Alternative program	39	2.8%
Not listed here	52	3.7%
Missing	68	4.8%

The teaching characteristics of survey respondents, including the subject and grades they teach, as well as their teaching experience are presented in Table 2. Most of our sample teacher ELA (40.8%) and Math (39.1%). Teacher in higher school grades, 9, 10, 11, 12 were 20.7%, 21.7%, 22.5% and 21.2%, respectively.

Table 2. Teaching Characteristics of the Participants

	N	%
Subject*		
ELA	579	40.8%
Math	554	39.1%
Science	167	11.8%
Social Studies	172	12.1%
Other	684	48.2%
Grade*		
Pre-K	201	14.2%
Grade 1	185	13.0%
Grade 2	172	12.1%
Grade 3	187	13.2%
Grade 4	204	14.4%
Grade 5	207	14.6%
Grade 6	210	14.8%
Grade 7	199	14.0%
Grade 8	201	14.2%
Grade 9	293	20.7%
Grade 10	308	21.7%
Grade 11	319	22.5%
Grade 12	301	21.2%
Experience		
I am a pre-service teacher	5	.4%
0-2 years	115	8.1%
3-5 years	170	12.0%
6-10 years	251	17.7%

More than 10 years	813	57.3%
Missing	64	4.5%

Note. *Check all that apply type items.

Measures

Five scales were used in our study, and they were teacher self-efficacy, teacher job satisfaction, teacher commitment, teacher collaboration, teacher-student interaction.

TSES

Prior research defines teacher self-efficacy as a measure of a teacher's judgment of their own ability to reach desired outcomes (Bandura, 1977). Tschannen-Moran and Hoy (2001) developed TSES, which asks teachers questions about how well they can perform various tasks within schools. Nie et al. (2012) examine the validity of TSES and suggested the scale with tree factors and 12 items in Singapore educational context. We used Nie et al.'s version of TSES in our study. The full list of teacher self-efficacy questions is listed in Table 3. For example, teachers were asked, *"How well can you respond to difficult questions from your students?"* and *"How well can you help your students value learning*?" All questions used five-point Likert type responses from *not well* at all to *very well*.

Teacher Job Satisfaction

Teacher job satisfaction was validated with a large international sample, including the United States, using confirmatory factor analysis (Pepe, 2011; Pepe et al., 2017). Thus, research provides evidence that this teacher job satisfaction scale is an appropriate tool to understanding teachers' level of job satisfaction. Teacher job satisfaction was measured as three sub-constructs: satisfaction with co-workers, students, and parents with nine items. For example, teachers were asked, "*How satisfied are you with the following aspect of the school: The extent to which your co-workers encourage you and support you in your work,"* and "*How satisfied with the following aspect of the school: The degree of interest shown by parents in the education of their children."*

Three-factor CFA analysis showed that the data reasonably fit this teacher job satisfaction. Specifically, the chi-square goodness-of-fit statistics are statistically significant, suggesting that the model fit is not perfect. However, the other goodness-of-fit statistics suggest that the data are a reasonable fit for the model (Chi-square = 117.84 (df=24), $p < .001$; CFI=.986; RMSEA=.053 [90% CI: .043 to .062]; SRMR=.033). Similar to Pepe et al.'s (2017) results, the standardized correlation coefficients between factors range from .29, .30 and .62 for students with co-workers, parents with co-workers and parents with students, respectively.

Teacher Commitment

Teacher commitment measures how dedicated teachers are to remaining in their profession. The teacher commitment measure was developed and validated by Thien et al. (2014). The researchers analyzed the results of over 600 teacher respondents using exploratory and confirmatory factor analysis. Their results provide evidence of construct validity for the teacher commitment scale. We used the sub-construct of commitment to the profession as an indication of teacher commitment. The teacher commitment to the profession construct contains four questions. Teachers were asked, for example, "*To what extent do you agree or disagree with the following statements: If I could get a job different from being a teacher and paying the same amount, I would take it.*" and "*To what extent do you agree or disagree with the following statements*: *One of the best decisions that I have ever made was to become a teacher."*

Analysis of from this survey administration provides evidence that the data reasonably fit this teacher commitment model. Since Chi-square goodness-of-fit statistics are statistically significant, the model fit is not perfect. However, the other goodness-of-fit statistics suggest that reasonable fit for the model (Chi-square = 7.67 (df=2), $p < .001$; CFI=.995; RMSEA=.045 [90% CI: .015 to .080]; SRMR=.009).

Teacher-Student Interaction

Brand et al. (2008) validated this scale as a part of school climate survey for teachers with a sample of 234 teachers. The researchers provide evidence for construct validity based on confirmatory factor analysis. The survey measures teacher-student interactions as one construct that includes five questions. These questions focus on interpersonal interactions. For example, teachers are asked, *"To what extent do you agree with the following statements: My students share their concerns with me,"* and "*To what extent do you agree with the following statements: My students express their feelings."*

Analysis from this survey administration provides evidence that the data reasonably fit this teacher-student interaction model. Specifically, the chi-square goodness-of-fit statistics are statistically significant, which suggestions that the model fit is not perfect. However, the other goodness-of-fit statistics suggest that a reasonable fit for the model (Chi-square = 27.52 (df=5), $p < .001$; CFI=.987; RMSEA=.056 [90% CI: .037 to .078]; SRMR=.019).

Data Analysis

We used SPSS 26 for all data analysis except for Confirmatory Factor Analysis (CFA) which was conducted on Mplus. Data was screened, and all assumptions, normality, outliers and adequate sample size were met. We used CFA for internal structure, correlation for convergent validity evidence, and Analysis of Variance (ANOVA) for criterion validity evidence. Lastly, Cronbach's Alpha and McDonald's Omega were used for the reliability.

CFA as Internal Structure Evidence

The internal structure represents to what degree the relationship between items and factors fits the construct (AERA, APA & NCME, 2014). We examined the validity evidence for TSES by conducting second-order CFA with 12 items as suggested by Nie et al. (2012). Before conducting CFA, the data was screened, and all assumptions were examined. Based on this analysis, multivariate outliers (N=66, 0.04%) were detected. CFA models were conducted with and without outliers. There was not a significant difference between the results, so outliers were not deleted. A second order CFA model with three-factor was conducted for TSES. All CFA models are overidentified, which indicates there is more than enough information in the data to estimate the model parameters. CFA models were tested with *M*plus 8 (Muthén & Muthén, 2017) using maximum likelihood estimation with robust standard errors. The first indicator of each latent variable' coefficient was fixed to 1.00.

We assessed model fit using the Chi-square test and along with the following goodness-of-fit indices: root mean square error of approximation (RMSEA); standardized root mean squared residual (SRMR); and comparative fit index (CFI). The Chi-square test assesses the difference between the given model and an unspecified model that would fit to the covariance matrix of the data perfectly (Kline, 2016, p. 270). While $p>.05$ is desired for Chi-square test, significant p-values may or may not indicate inappropriate model fit in large sample studies. Thus, we used other indices to test how well the model fit the data. RMSEA is a based-on error terms; thus, zero is the best result (Kline, 2016, *p*.273). For RMSEA, values greater than .10 may indicate a lack of fit (Browne & Cudeck, 1992). CFI is a goodness-of-fit indices, and CFI values greater than .90 indicates that the proposed model is greater than 90% of than that of the baseline model, serve as an indicator of adequate fit (Kline, 2016). SRMR is standardized measure of the absolute covariance residual, and perfect model fit is indicated by SRMR = 0, and values greater than .10 may indicate poor fit (Kline, 2016).

Pearson Correlation as Convergent Validity Evidence

Examining the relationship between the construct and other related variables serves convergent validity evidence (AERA, APA, & NCME, 2014). We used Pearson correlation to analyze the relationship between overall teacher self-efficacy, three subscales of TSES, teacher job satisfaction, teacher commitment and teacher-student interaction. The overall teacher self-efficacy, three subscales of TSES, teacher job satisfaction, teacher commitment and teacher-student interaction variables were computed as the mean of all items in the scale.

ANOVA as Criterion Validity Evidence

Criterion validity examines the relationship between the target construct and a relevant criterion (AERA, APA, & NCME, 2014). Literature demonstrates a positive correlation between teacher self-efficacy and years of teaching experience (Gale et al., 2021). Therefore, we used years of teaching experience as a criterion. We grouped years of teaching experience into three groups, 0-5 years, 6-10 years and more than 10 years to conduct

a one ANOVA because we expected teacher with more experience to have a higher teaching self-efficacy.

Reliability

We used Cronbach's Alpha and McDonald's Omega to provide reliability evidence. Cronbach's Alpha is a reliability index and a measure of internal consistency. Cronbach's Alpha shows the relationship between question responses in the same scale. A higher Cronbach's Alpha indicates a higher reliability of the scale. If Cronbach's Alpha is higher than the .70 for a group of questions, then those questions have an acceptable reliability index (Nunnally, 1978). Well-accepted value for McDonald's omega is .9 (Tervee et al., 2007).

Results

Our study demonstrates the three types of validity evidence for TSES, internal structure validity, convergent validity and criterion validity, and reliability evidence.

Internal Structure (CFA)

Table 3. Descriptive Statistics for Teacher Self-efficacy Items and Constructs

Item ID	Items and Constructs	Mean	SD
	Instructional strategies (IS)	**3.97**	**0.58**
TS1	How well can you respond to difficult questions from your students?	4.02	0.72
TS2	How well can you provide appropriate challenges for very capable students?	3.77	0.79
TS3	How well can you implement alternative instructional strategies in your classroom?	3.91	0.79
TS4	How well can you provide an alternative explanation, for example, when students are confused?	4.17	0.70
	Motivation (MOT)	**3.50**	**0.73**
TS5	How well can you help your students value learning?	3.68	0.83
TS6	How well can you motivate students who show low interest in schoolwork?	3.32	0.91
TS7	How well can you improve the understanding of a student who is failing?	3.57	0.81
TS8	How well can you get through to the most difficult students?	3.41	0.92
	Classroom management (CM)	**4.01**	**0.69**
TS9	How well can you make your expectations clear about student behavior?	4.34	0.71
TS10	How well can you get students to follow classroom rules?	4.07	0.77
TS11	How well can you control disruptive behavior in the classroom?	3.99	0.82
TS12	How well can you keep a few problem students from missing an entire lesson?	3.66	0.92
	Teacher self-efficacy (TS)	**3.83**	**0.56**

Our results showed that the data fits the model well based on the goodness-of-fit statistics. The chi-square goodness-of-fit statistics were statistically significant, suggesting the model fit is not perfect. However, all other goodness-of-fit statistics provide evidence that the data do adequately fit the model (Chi-square = 450.33 (df=51), $p < .001$; CFI=.941; RMSEA=.074 [90% CI: .068 to .081]; SRMR=.042). The unstandardized and standardized coefficients for teacher self-efficacy second-order CFA are reported in Table 4. All coefficients were statistically significant ($p<.001$) indicating that the coefficients are larger than zero. The squared valued of standardized coefficient shows the proportion of explained variance. Therefore, any standardized coefficient that falls below .70 indicates that less than half of the variation in that question response is accounted for in the factor. Thus, Table 4 shows that there are three questions with standardized coefficients below .70, within the Instructional Strategies factor. Although these values are below .7 threshold, they are not far from .7. For example, the standardized coefficient of TS1 was .65, indicating that this question explains 42% of variance in its factor. Figure 1 contains a visual representation of this model.

Table 4.Unstandardized Coefficients, Standard Error (SE), and Standardized Coefficients for Teacher Self-efficacy CFA Model

Item ID	Constructs and Questions	Unstandardized		Standardized
		Coefficient	SE	Coefficient
	Instructional strategies (IS)			
TS1	How well can you respond to difficult questions from your students?	1.00	0.00	0.65
TS2	How well can you provide appropriate challenges for very capable students?	1.21	0.06	0.71
TS3	How well can you implement alternative instructional strategies in your classroom?	1.18	0.07	0.69
TS4	How well can you provide an alternative explanation, for example, when students are confused?	1.01	0.05	0.67
	Motivation (MOT)			
TS5	How well can you help your students value learning?	1.00	0.00	0.78
TS6	How well can you motivate students who show low interest in schoolwork?	1.18	0.03	0.84
TS7	How well can you improve the understanding of a student who is failing?	0.90	0.04	0.72
TS8	How well can you get through to the most difficult students?	1.09	0.04	0.77
	Classroom management (CM)			
TS9	How well can you make your expectations clear about student behavior?	1.00	0.00	0.74
TS10	How well can you get students to follow classroom rules?	1.27	0.05	0.87
TS11	How well can you control disruptive behavior in the	1.37	0.05	0.88

	classroom?			
TS12	How well can you keep a few problem students from missing an entire lesson?	1.31	0.06	0.75
	Teacher Self-efficacy			
	IS	1.00	0.00	0.85
	MOT	1.38	0.09	0.84
	CM	0.99	0.06	0.74

Note. IS=Instructional strategies, MOT= Motivation, CM= Classroom management

The path coefficients between the factors and second order factor are statistically significant at .85, .84 and .74 for instructional strategies, motivation, and classroom management, respectively. These factors explain 72%, 70% and 55% of the variance in teacher self-efficacy, respectively. The teacher self-efficacy second-order CFA model is shown in Figure 1. This figure provides a visual representation of the model. That is, the figure shows how the teacher self-efficacy (ts) construct, relates to the three sub-constructs of instructional strategies (is), motivation (mot), and classroom management (cm). Similarly, the figure shows how each question item relates to subconstructs and the overall construct.

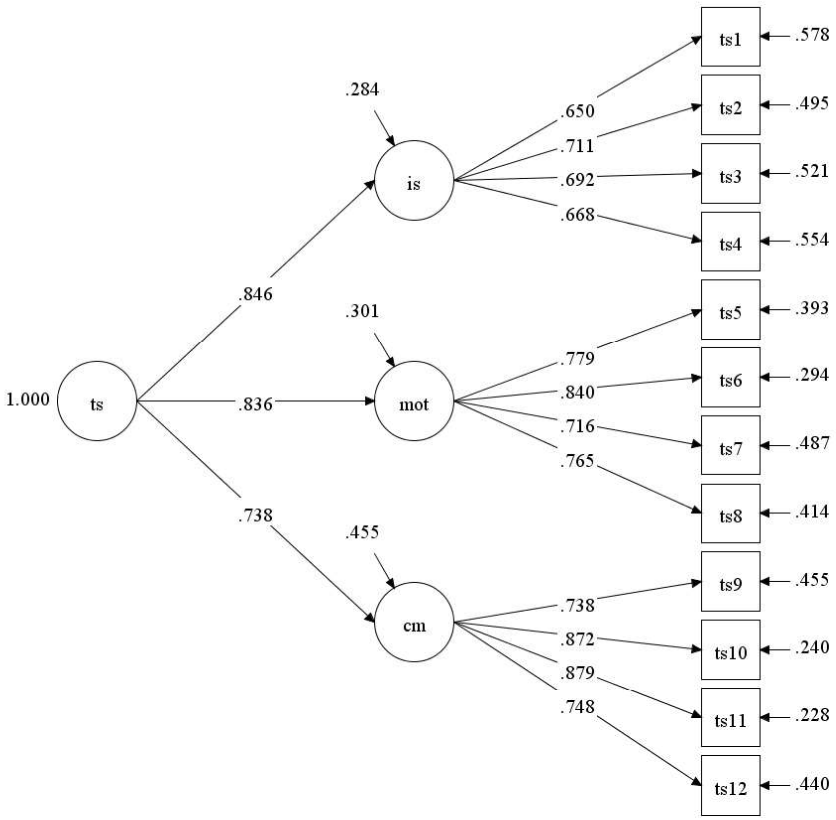

Figure 1. Teacher Self-efficacy Second-order CFA Model with Standardized Estimations

Note. is=Instructional strategies, mot= Motivation, cm= Classroom management, ts=Teacher self-efficacy.

Pearson Correlation as Convergent Validity Evidence

Table 5 displays the Pearson correlation among TSE, TSE subscales, teacher job satisfaction, and teacher student interaction. Our results showed that three subscales of TSE were highly correlated ($p< .01$). Besides, TSE is correlated with teacher job satisfaction and teacher student interaction.

Table 5. Correlation Matrix among TSE and other variables

Variables	1	2	3	4	5
1- TSE					
2-TSE-Instructional strategies	.818**				
3-TSE-Motivation	.867**	.589**			
4-TSE-Classroom management	.838**	.535**	.565**		
5- Teacher job satisfaction	.274**	.116**	.306**	.248**	
6- Teacher-student interaction	.334**	.246**	.381**	.205**	.249**

Note. *p<.05; **p<.01; ***p<.001

ANOVA as Criterion Validity Evidence

A one-way ANOVA was conducted to determine if there is mean difference in TSE based on teacher years of experience. According to Levene's test, there was heterogeneity among three groups, so we used Welch test (Lomax & Hahs-Vaughn, 2013). Our results indicated that there is a significant difference in TSE based on teacher years of experience ($F(2, 508.39) = 51.85$, $p<.001$, $\eta^2=.08$). The posthoc test indicated that there was a significant difference in TSE for 0-5 years of experience (M=3.54, SD= .6), 6-10 years of experience (M=3.82, SD=.55) and more than 10 years of experience (M=3.93, SD=.52) which indicated that more experience teachers have higher TSE.

Survey Reliability

Cronbach's Alpha, a common measure of reliability, is calculated for each scale and sub-scales. Table 6 presents the number of items and item reliability index, Cronbach's Alpha and Mc Omega for TSE and its subscales. TSE presented higher reliability based on Cronbach' Alpha (.901) and Mc Omega (.898). Therefore, we concluded that TSE scale has high reliability for this sample.

Table 6. The Item Analysis of the Scales

Scale	Subscales	Number of items	Cronbach's Alpha	Mc Omega
Teacher Self-efficacy		12	.901	.898

	Instructional Strategies	4	.773	.77
	Motivation	4	.855	.857
	Classroom Management	4	.876	.88

Conclusion

This study examined the validity and reliability evidence of TSE (TschannenMoran & Hoy, 2001) scale with a large US sample. CFA results provided construct validity evidence for TSE as suggested by (Nie et al., 2012) for teacher in Singapore. These results indicate that the scale can be used to measure teacher self-efficacy in the United States.

Our study presented convergent validity evidence and high reliability for TSE. Three factors of TSE, instructional strategies, motivation and classroom management, were highly correlated with overall TSE, and moderately correlated with teacher job satisfaction, and teacher-student interaction. These positive correlations between TSE and teacher job satisfaction and teacher-student interaction served as convergent validity evidence of TSE. The ANOVA results presented that there was a significant difference in TSE based on teaching experience as more experience teacher had higher level of TSE (Gale et al., 2021; George et al., 2018; Wolters & Daugherty, 2007). This result supported criterion validity evidence for TSE. Based on Cronbach's Alpha and Mc Omega TSE and its three factors have high reliability indices.

References

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. American Educational Research Association.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84, 191–215.

Bloom, J. L., & Peters, S. J. (2012). Lessons learned from urban teachers' self-efficacy reports: Insights into building a supportive school culture. *Urban Education*, 47(1), 77-107.

Brand, S., Felner, R. D., Seitsinger, A., Burns, A., & Bolton, N. (2008). A large scale study of the assessment of the social environment of middle and secondary schools: The validity and utility of teachers' ratings of school climate, cultural pluralism, and safety problems for understanding school effects and school improvement. *Journal of School Psychology, 46*(5), 507-535. doi:10.1016/j.jsp.2007.12.001

Cano-García, F. J., Padilla-Muñoz, E. M., & Carrasco-Ortiz, M. Á. (2005). Personality and contextual variables in teacher burnout. *Personality and Individual Differences, 38*, 929-940. doi:10.1016/j.paid.2004.06.018

DeVilles, R. F. (2012). *Scale development: Theory and applications*. Sage Publications.

Fathi, J., & Rostami, M. (2018). Investigating the psychometric properties of teacher self-efficacy scale in Iran. *Educational Research and Evaluation*, 24(5-6), 312-330.

Gale, J., Alemdar, M., Cappelli, C., & Morris, D. (2021). A mixed methods study of self-efficacy, the sources of self-efficacy, and teaching experience. *Frontiers in Education*. https://doi.org/10.3389/feduc.2021.750599

Gavish, B., & Friedman, I. A. (2010). Novice teacher's experience of teaching: A dynamic aspect of burnout. *Social Psychology of Education, 13*, 141-167. doi:10.1007/s11218-009-9108-0

George, S. V., Richardson, P. W., and Watt, H. M. (2018). Early career teachers' self-efficacy: A longitudinal study from Australia. *Aust. J. Edu*. 62 (2), 217–233. doi:10.1177/0004944118779601

Kasalak, K. J., & Dagyar, M. (2020). The relationship between job satisfaction and teacher self-efficacy: A meta-analysis study. *Educational Sciences: Theory & Practice*, 20(1), 145-161.

Klassen, R. M., Bong, M., Usher, E. L., Chong, W. H., Huan, V. S., Wong, I. Y. F., & Georgiou, T. (2009). Exploring the validity of a teachers' self-efficacy scale in five countries. *Contemporary Educational Psychology*, 34(1), 67-76.

Kline, R. B. (2016). *Principles and practice of structural equation modeling*. Guilford publications.

Lomax, R. G., & Hahs-Vaughn, D. L. (2013). *An introduction to statistical concepts*. Routledge.

Ma, K., Trevethan, R., & Lu, S. (2019). Measuring teacher sense of efficacy: Insights and recommendations concerning scale design and data analysis from research with preservice and inservice teachers in China. *Frontiers of Education in China, 14*, 612–686. https://doi.org/10.1007/s11516-020-0009-5.

Muthén, B., & Muthén, L. (2017). *Mplus. In Handbook of item response theory* (pp. 507-518). Chapman and Hall/CRC.

Nie, Y., Lau, S., & Liau, A. (2012). The teacher efficacy scale: A reliability and validity study. *The Asia-Pacific Education Researcher 21*(2), 414-421.

Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York: McGraw-Hill.

Pepe, A. (2011). Measuring teacher job satisfaction: A quantitative empirical tool. Paper presented at the *8th International Conference of European Research Network About Parents in Education*, Milano, Italy.

Pepe, A., Addimando, L., & Veronese, G. (2017). Measuring teacher job satisfaction: Assessing invariance in the teacher job satisfaction scale (TJSS) across six countries. *Europe's journal of psychology, 13*(3), 396.

Siwatu, K. O., & Starker, T. V. (2010). Predicting preservice teachers' self-efficacy to resolve a cultural conflict involving an African American student. *Multicultural*

Perspectives, 12(1), 10–17. https://doi.org/10.1080/15210961003641302.

Skaalvik, E. M., & Skaalvik, S. (2011). Teachers' feeling of belonging, exhaustion, and job satisfaction: The role of school goal structure and value consonance. *Anxiety, Stress, and Coping, 24*(4), 369-385. doi:10.1080/10615806.2010.544300

Summers, J. J., Davis, H. A., & Woolfolk Hoy, A. (2017). The effects of teachers' efficacy beliefs on students' perceptions of teacher relationship quality. *Learning and Individual Differences*, 53, 17–25. https://doi.org/10.1016/j.lindif.2016.10.004.

Sylvia, R. D., & Hutchinson, T. (1985). What makes Ms. Johnson teach? A study of teacher motivation. *Human Relations*, 38, 841-856. doi:10.1177/001872678503800902

Terwee, C.B.; Bot, S.D.M.; de Boer, M.R.; van der Windt, D.A.W.M.; Knol, D.L.; Dekker, J.; Bouter, L.M.; de Vet, H.C.W. Quality criteria were proposed for measurement properties of health status questionnaires. *J. Clin. Epidemiol*. 2007, 60, 34–42.

Thien, L. M., Razak, N. A., & Ramayah, T. (2014). Validating teacher commitment scale using a Malaysian sample. *Sage Open, 4*(2). doi:10.1177/2158244014536744

Tschannen-Moran, M., & Hoy, A. W. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and Teacher Education, 17*(1), 783-805. doi: 10.1016/S0742- 051X(01)00036-1

Van Droogenbroeck, F., Spruyt, B., & Vanroelen, C. (2014). Burnout among senior teachers: Investigating the role of workload and interpersonal relationships at work. *Teaching and Teacher Education*, *43*, 99-109. doi:10.1016/j.tate.2014.07.005

Wolters, C. A., and Daugherty, S. G. (2007). Goal structures and teachers' sense of efficacy: Their relation and association to teaching experience and academic level. *J. Educ. Psychol*. 99 (1), 181–193. doi:10.1037/0022-0663.99.1.181

Worrell, T. G., Skaggs, G. E., & Brown, M. B. (2006). School psychologists' job satisfaction: A 22-year perspective in the USA. *School Psychology International, 27*(2), 131-145. doi:10.117