INTERNATIONAL JOURNAL OF ACADEMIC RESEARCH IN PROGRESSIVE EDUCATION & DEVELOPMENT

OPEN ACCESS JOURNAL

Vol 13, Issue 4, (2024) E-ISSN: 2226-6348

Validity and Reliability of the Chinese Version of the Utrecht Work Engagement Scale for Students in a Sample of Early Childhood Pre-Service Teachers

Sha Xinchong[1,2], Mohd Mokhtar Muhamad[1], Fazilah Razali[1], Nasnoor Juzaily Mohd Nasiruddin[3], Shao Mingxing[1,2], Lai Zhujing[2]

[1]Faculty of Educational Studies, Universiti Putra Malaysia, Selangor, Malaysia, [2]Faculty of Preschool Education, Qiongtai Normal University, Haikou, China, [3]Faculty of Sports and Exercise Science, University of Malaya, Kuala Lumpur, Malaysia

Corresponding Author Email: qdshaomingxing@gmail.com

To Link this Article: http://dx.doi.org/10.6007/IJARPED/v13-i4/23604 DOI:10.6007/IJARPED/v13-i4/23604

Published Online: 26 December 2024

Abstract

The purpose of this study was to test the validity and reliability of the Utrecht Work Engagement Scale for Students (UWES-S) among Chinese early childhood pre-service teachers (ECPTs) and to corroborate the factor structure of both the UWES-17S and UWES-9S. The study examined the UWES-17S and UWES-9S through CFA, convergent validity, reliability, and inter-correlations with a sample of 515 ECPTs in China. The modified UWES-17S and UWES-9S models showed acceptable fit indices, with the 9-item, 3-factor model showing a slightly better fit than the 17-item version. Both models had good convergent validity, with the UWES-9S performing slightly better. In both versions, strong positive correlations were found across all dimensions. The UWES-17S showed slightly higher internal consistency across the total scale and its three factors. The study confirmed the 3-factor structure of UWES-17S and UWES-9S, confirming that they are reliable tools for assessing Chinese ECPTs' learning engagement. The UWES-9S model is simpler and has a better fit, while the UWES-17S has an advantage in reliability.

Keywords: UWES-S, ECPTs, Validity and Reliability, China

Introduction

Engagement is described by Schaufeli et al. (2002), as a positive, fulfilling, work-related mental state. As an aspect of positive psychology, engagement not only enhances performance and outcomes but also effectively reduces negative psychological states and symptoms, such as burnout (Tatha et al., 2024). For students, learning engagement represents a sustained state and positive emotions throughout the learning process, typically characterized by vigor, dedication, and absorption (Schaufeli et al., 2002; Qi et al., 2024). This

concept is closely associated with students' learning status and serves as a crucial indicator for assessing early childhood teacher education programs (Qin et al., 2021).

Studying learning engagement is crucial for ECPTs. Previous studies have shown that learning engagement is closely related to academic achievement, psychological health, and learning burnout, and lack of engagement can lead to poor academic performance or certain psychological problems (Qi et al., 2024). Improving learning engagement can not only improve ECPTs' self-efficacy, but also help reduce their learning burnout and stress, and improve student performance (Yang, 2017; Qin et al., 2021). In the context of Chinese teacher education, focusing on improving learning engagement and effectively evaluating it can significantly improve the quality and impact of training programs (Qin et al., 2021). Therefore, using scientific tools to measure the learning engagement of ECPTs is crucial to understanding the learning engagement of this group. These tools can help educators, scholars, and policymakers identify factors that affect engagement, support teacher training, guide curriculum design, and ultimately promoting the professional development of early childhood educators and improving early childhood education.

Schaufeli et al. (2002) developed the UWES-S to measure students' academic engagement. The original UWES-S included 17 items, but the team later published a shortened version with 9 items (Schaufeli et al., 2006). For the 17-item UWES-S, Schaufeli et al. (2002) discovered that the 3-factor model fit the data better than the 1-factor model among 314 undergraduates in Spain. Similarly, Fang et al. (2008) reported good reliability and validity for the scale in Chinese undergraduate and graduate samples, confirming the 3-factor structure. These findings were consistent with Meng and Jin's (2017) study on Chinese nursing college students. In the Korean context, Jang and An (2022) also found the 17-item version to be a reliable instrument with an applicable 3-factor structure. Likewise, the validity and reliability of the 9-item UWES-S have been supported by several studies (Carmona-Halty et al., 2019; Chi et al., 2023).

However, previous research results have shown some inconsistencies. For example, Schaufeli and Bakker (2004) observed that the model fit of the 9-item scale was better than that of the 17-item scale in the context of 10 different Western countries, with the 3-factor structure of the 9-item scale outperforming the 1-factor model. In contrast, Chi et al. (2003) found that the 1-factor, 9-item scale had the best goodness-of-fit index in the Chinese context. Additionally, the existing body of research on this topic is relatively limited, and the reliability and validity of the UWES-S for Chinese ECPTs remain uncertain.

This study aims to address these gaps by testing the reliability and validity of the UWES-S among pre-service teachers in early childhood education in China and confirming the factor structure of the scale. Our goal is to provide a scientifically robust and accurate research tool, thereby contributing to the practice and research in early childhood pre-service education among Chinese pre-service teachers.

Methods and Materials

Methodology and Sample

This study is a cross-sectional validation study designed to evaluate the reliability and validity of the UWES-S among ECPTs in China. After receiving approval from the Universiti Putra

Malaysia Ethics Committee (JKEUPM-2023-339), we randomly distributed online questionnaires to ECPTs at four universities in China. A total of 515 ECPTs consented to participate and completed the questionnaire online. They were informed that all personal information would be used only for research purposes and would be securely protected.

Instrument

This study employed the UWES-S, developed by Schaufeli et al. (2002). We evaluated the reliability and validity of two versions of the scale: the 17-item UWES-S (UWES-17S) and the 9-item short form (UWES-9S), which excludes items VI1, VI2, VI6, DE1, DE5, AB1, AB2, and AB6. The UWES-17S includes a 3-factor structure comprising vitality (VI), dedication (DE), and absorption (AB). For the UWES-9S, we evaluated both the 3-factor and 1-factor structures. All items were rated on a 7-point Likert scale.

Data Analysis

In this study, data analysis was performed using Amos version 24 and SPSS version 22. First, a preliminary analysis was performed to calculate the mean, standard deviation (SD), skewness, and kurtosis. Skewness and kurtosis values between -1.00 and +1.00 were considered highly normally distributed (George & Mallery, 2003). Next, we used CFA to evaluate three initial models and two modified models with the following criteria: $\chi^2/df \leq 3$ (Kline, 1998), GFI≥0.90 (Kline, 2005), AGFI≥0.90 (Tabachnick & Fidell, 2007), RMSEA≤0.05 (MacCallum et al., 1996), TLI (NNFI)>0.90 (Byrne, 1994), CFI≥0.90 (Fan et al., 1999), IFI≥0.90 (Bollen, 1989). The lower the AIC value, the better the fit. Then, we assessed the convergent validity of the modified UWES-17S and UWES-9S with an AVE threshold of 0.50 and a CR threshold of 0.60 (Bagozzi & Yi, 1988). Reliability was measured using Cronbach's Alpha, and values above 0.70 were considered acceptable (Karagöz, 2018). Finally, Pearson's two-tailed correlation was used to analyze the relationship between the sub-scales and the total scale of UWES-17S and UWES-9S. When r exceeds 0.70, the correlation is considered very strong (Cohen, 1988).

Findings

Preliminary Analysis

Table 1

Descriptive statistics of UWES-S

Items	Sample (n=515)			
	Mean	SD	Skewness	Kurtosis
VI1	4.58	1.44	-0.35	0.10
VI2	4.24	1.32	0.03	0.24
VI3	4.49	1.36	-0.17	0.14
VI4	4.49	1.36	-0.04	-0.01
VI5	3.82	1.52	0.25	-0.24
VI6	4.53	1.31	0.03	0.06
DE1	5.02	1.39	-0.29	-0.16
DE2	4.74	1.37	-0.19	-0.02
DE3	4.45	1.36	0.04	0.10
DE4	4.89	1.38	-0.21	-0.15
DE5	5.18	1.29	-0.25	-0.11
AB1	4.86	1.42	-0.26	-0.09
AB2	4.64	1.36	-0.01	-0.13
AB3	4.88	1.38	-0.23	-0.04
AB4	4.62	1.32	-0.02	0.14
AB5	4.63	1.34	-0.04	0.04
AB6	4.29	1.53	0.05	-0.26

The descriptive statistics in Table 1 provide insights into the learning engagement levels of ECPTs as assessed by the 17-item UWES-S, which includes all items on the UWES-9S. Based on the table, the mean scores ranged from approximately 3.82 to 5.18, with standard deviations from 1.29 to 1.53, indicating that ECPTs exhibited a moderately high level of learning engagement, albeit with some variations within the sample. The skewness and kurtosis values for all items fall within the range of ±1, typically considered indicative of a normal distribution (George & Mallery, 2003). Therefore, these data were appropriate for further analysis to assess the psychometric properties of the two versions of UWES-S.

CFA

Table 2

The CFA Outcomes of the Chinese Version of UWES-S (n=515)

	χ^2	df	χ^2/df	GFI	AGFI	RMSEA	CFI	TLI (NNFI)	IFI	AIC
UWES-17S (3 factor)	704.50	116	6.07	0.85	0.80	0.99	0.93	0.92	0.93	778.50
UWES-9S (3 factor)	191.83	24	7.99	0.93	0.86	0.12	0.96	0.94	0.96	233.83
UWES-9S (1 factor)	408.09	27	15.11	0.85	0.74	0.17	0.92	0.89	0.92	444.09
Modified UWES-17S (3 factor)	229.21	116	1.98	0.97	0.96	0.04	0.99	0.98	0.99	303.21
Modified UWES-9S (3 factor)	55.44	24	2.31	0.99	0.90	0.05	0.99	0.99	0.99	97.44

Compared with the UWES-17S (3-factor) model, the UWES-9S (3-factor) model demonstrated better-fit indices, including GFI (0.93), CFI (0.96), TLI (0.94), and IFI (0.96). Moreover, its AIC (233.83) was lower, indicating a more parsimonious model. However, the χ^2/df (>3), AGFI (<0.90), and RMSEA (>0.05) results in both models, as well as GFI in the UWES-17S (3-factor) model (<0.90) exhibited a poor fit. Additionally, the results showed that the UWES-9S (1-factor) model presented a worse fit than the UWES-9S (3-factor) model and UWES-17S (3-factor) model.

Then, we applied the Bollen bootstrap method to modify the UWES-17S (3-factor) and UWES-9S (3-factor) models. The fit indices of the two modified models reached the acceptable threshold, with the 9-item 3-factor model fitting slightly better than the 17-item 3-factor model.

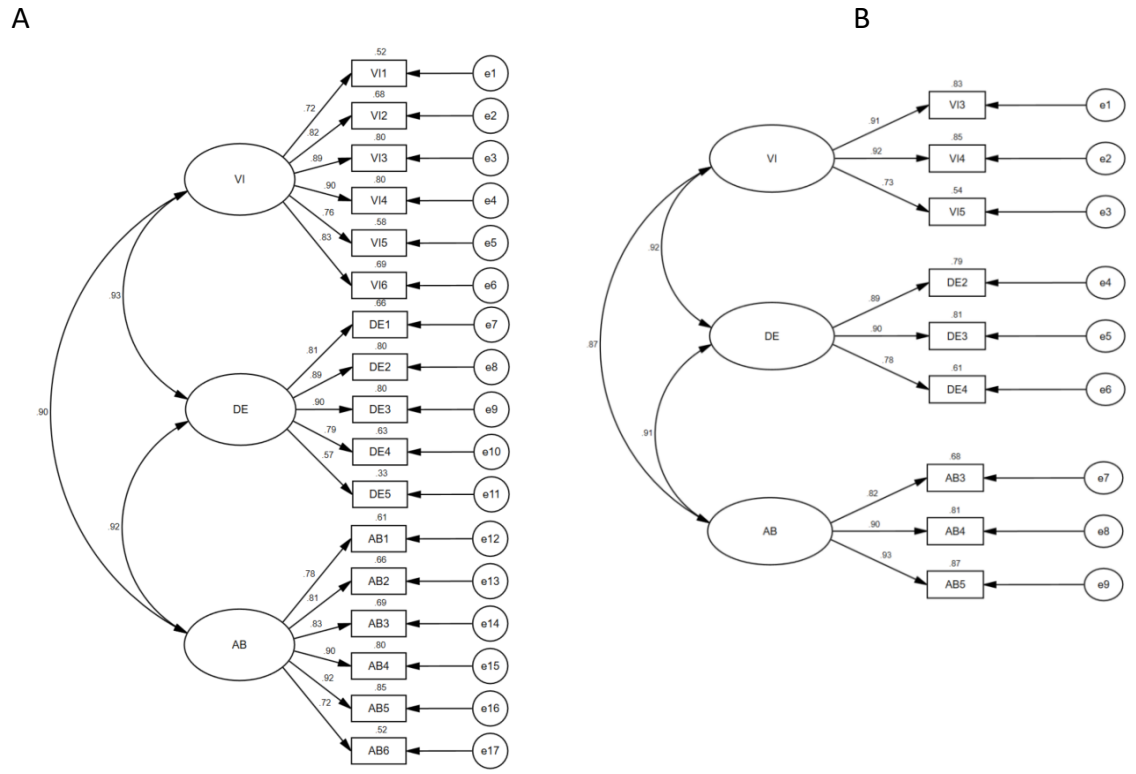

Figure 1. The factor structure of UWES-S (A: 3 factor UWES-17S, B: 3 factor UWES-9S)

Convergent Validity

Table 3

AVE and CR of the UWES-S (n=515)

Variable	UWES-17S		UWES-9S	
	AVE	CR	AVE	CR
VI	0.68	0.93	0.74	0.89
DE	0.64	0.90	0.74	0.89
AB	0.69	0.93	0.78	0.92

Table 3 presents the AVE and CR values for the UWES-S across different model configurations (n=515). The results indicated that both 3-factor models were robust and reliable. Notably, the AVE values of the UWES-9S were slightly higher than those of the UWES-17S, indicating that the latent variables in the UWES-9S could better capture the variance of the observed variables. Thus, we believed that the UWES-9S had advantages in simplicity and measurement accuracy.

Reliability

Table 4

Cronbach's alpha coefficients of the UWES-S (n=515)

Variable	UWES-17S	UWES-9S
VI	0.92	0.88
DE	0.90	0.89
AB	0.92	0.91
UWES-S	0.97	0.95

As shown in Table 4, both the UWES-17S and UWES-9S models demonstrated high internal consistency across all dimensions, with α values significantly exceeding 0.70. The UWES-17S model exhibited slightly higher internal consistency across the three factors compared to the UWES-9S. Overall, both the UWES-17S and UWES-9S models were reliable for measuring the learning engagement of Chinese ECPTs within a 3-factor structure.

Inter-correlations

Table 5

Correlations between Sub-Scales and the UWES-S (n=515)

Variable	Total scale	VI	DE	AB
Total scale	1	0.94***	0.95***	0.94***
VI	0.94***	1	0.84***	0.80***
DE	0.94***	0.82***	1	0.85***
AB	0.96***	0.85***	0.86***	1

*Note. Below diagonal, UWES-17S; above diagonal, UWES-9S; *** $p \leq 0.001$.*

Table 5 displays the correlations between the sub-scales of the UWES-S and the total scale, as well as between the sub-scales. For the UWES-17S, the correlation coefficients between the sub-scales (VI, DE, and AB) and the total scale ranged from 0.94 to 0.96, while in the UWES-9S, the correlation coefficients ranged between 0.94 and 0.95. In addition, the correlation coefficients among all dimensions in both versions of the scale showed strong positive correlations. The above results illustrated that both UWES-17S and UWES-9S were effective in measuring learning engagement, although the strength of these relationships differed slightly.

Discussion

The present study aims to assess the factor structure, reliability, and validity of the Chinese version of UWES-S among ECPTs. Our study fills the gap in theoretical research on the UWES-17S and UWES-9S in the context of ECPTs in China and provides a solid theoretical foundation for future research.

The initial CFA results indicated that the 3-factor UWES-9S model fitted better, followed by the 3-factor UWES-17S model, except for the χ^2/df, AGFI, and RMSEA metrics that were not fitted. After making some modifications to these two models by the Bollen bootstrap method, both achieved satisfactory fit indices, confirming the 3-factor structure of the UWES-S in this study. This finding is consistent with previous research on the UWES-17S conducted in Spain (Schaufeli et al., 2002), China (Meng & Jin, 2017), and Korea (Jang & An, 2022). Similarly, our finding of the 3-factor structure of the UWES-9S is consistent with a study from Chile (Carmona-Halty et al., 2019). However, a study using the UWES-9S found that the 1-factor structure provided the best fit in the context of Chinese in Taiwan (Chi et al., 2023). These inconsistent findings may be due to differences in the sample's age or professional background (Willmer et al., 2019). Therefore, further research is needed to examine the scale's applicability across different cultural and professional contexts.

In addition, we found that both the UWES-17S and UWES-9S demonstrated adequate convergent validity. The UWES-9S showed a slightly higher AVE value compared to the UWES-17S, which suggests it might be more effective in capturing and measuring observed variables.

Both versions also exhibited high reliability, with the UWES-17S showing slightly higher internal consistency across the total scale and its three factors. Moreover, there were strong positive correlations between the total scale and sub-scales, as well as among the sub-scales themselves, in both versions. The outcomes in the present study are consistent with Schaefer's findings (Schaufeli et al., 2002). These inter-correlations confirm that both scales are effective for measuring learning engagement.

In conclusion, our findings verify the 3-factor structure of both versions of the UWES-S scale, indicating that the UWES-17S and UWES-9S are reliable tools for assessing Chinese ECPTs' learning engagement. The UWES-9S model offers greater simplicity and better fit, which provides a more scientific and practical tool for assessing pre-service teachers' engagement in learning, while the UWES-17S has advantages in terms of reliability. Depending on the needs of the research, scholars can choose the more concise 9-item scale for better fit, or the 17-item scale for enhanced reliability.

Limitation and Recommendation

One limitation of this study is its cross-sectional design. Future research could adopt longitudinal research to investigate the changes in learning engagement over time and identify its influencing factors. In addition, exploring the measurement invariance of the UWES-S across different samples (e.g., by age or gender) would provide further insights into the applicability of the scale.

Acknowledgment

This study was supported by the projects Hainan Province's "South China Sea Rising Star" Education Platform Project (JYNHXX2023-24T), Educational and Teaching Reform of Hainan's Higher Education Institutions Research Project (Hnjgzc2022-67; Hnjg2024-155), and Qiongtai Normal University Research Grant (qtjg2024-07).

References

Bagozzi, R., & Yi, Y. (1988). On the Evaluation of Structural Equation Models. Journal of the Academy of Marketing Sciences, 16, 74-94. https://doi.org/10.1007/BF02723327

Bollen, K. A. (1989). Structural equations with latent variables. John Wiley & Sons. https://doi.org/10.1002/9781118619179

Byrne, B. M. (1994). Structural Equation Modelling with EQS and EQS/Windows: Basic Concepts, Applications, and Programming. Sage.

Carmona-Halty, M. A., Schaufeli, W. B., & Salanova, M. (2019). The Utrecht Work Engagement Scale for Students (UWES-9S): Factorial Validity, Reliability, and Measurement Invariance in a Chilean Sample of Undergraduate University Students. Frontiers in Psychology, 10, 1017. https://doi.org/10.3389/fpsyg.2019.01017

Chi, L.-C., Tang, T.-C., & Tang, E. (2023). Psychometric properties of the Utrecht Work Engagement Scale for Students (UWES-S) in the Taiwanese context. Current Psychology, 42(31), 27428–27441. https://doi.org/10.1007/s12144-022-03737-0

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Fan, X. T., Thompson, B., & Wang, L. (1999). Effects of sample size, estimation methods, and model specification on structural equation modeling fit indexes. Structural Equation Modeling: A Multidisciplinary Journal, 6(1), 56–83. https://doi.org/10.1080/10705519909540119

Fang, L. T., Shi, K., & Zhang, F. H. (2008). Reliability and validity of the Chinese version of the Learning Engagement Scale. Chinese Journal of Clinical Psychology, 16(6), 618-620. https://doi.org/10.16128/j.cnki.1005-3611.2008.06.023

George, D., & Mallery, P. (2003). SPSS for Windows Step by Step: A simple guide and reference, 11.0 (3rd ed.). USA: Allyn & Bacon.

Jang, A., & An, M. (2022). Korean Version of the 17-Item Utrecht Work Engagement Scale for University Students: A Validity and Reliability Study. Healthcare, 10(4), 642. https://doi.org/10.3390/healthcare10040642

Karagöz, Y. (2018). SPSS ve AMOS 23 applied statistical analysis (1st ed.). Ankara: Nobel.

Kline, R. B. (1998). Principles and practice of structural equation modeling. Guilford Press.

Kline, R. B. (2005). Principles and practice of structural equation modeling (2nd ed.). Guilford Press.

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. Psychological Methods, 1(2), 130–149. https://doi.org/10.1037/1082-989X.1.2.130

Meng, L., & Jin, Y. (2017). A confirmatory factor analysis of the Utrecht Work Engagement Scale for Students in a Chinese sample. Nurse Education Today, 49, 129-134. https://doi.org/10.1016/j.nedt.2016.11.017

Qi, X., Jiang, Y., & Lian, R. (2024). The effect of social media upward comparison on Chinese adolescent learning engagement: A moderated multiple mediation model. BMC Psychology, 12(1), 122. https://doi.org/10.1186/s40359-024-01621-z

Qin, X. X., Cheng, X. H., & Dong, W. (2021). The influence of professional identity on the academic self-efficacy of early childhood education students. Renmin University of China Journal of Education, (2), 121-135.

Schaufeli, W. B., & Bakker, A. B. (2004). Utrecht Work Engagement Scale: Preliminary manual (Version 1.1). Occupational Health Psychology Unit, Utrecht University. Retrieved from

https://www.wilmarschaufeli.nl/publications/Schaufeli/Test%20Manuals/Test_manual_UWES_English.pdf

Schaufeli, W. B., Salanova, M., González-Romá, V., & Bakker, A. B. (2002). The measurement of engagement and burnout: A two sample confirmatory factor analytic approach. Journal of Happiness Studies: An Interdisciplinary Forum on Subjective Well-Being, 3(1), 71–92. https://doi.org/10.1023/A:1015630930326

Schaufeli, W.B., Bakker, A.B., Salanova, M., (2006). The measurement of work engagement with a short questionnaire. A cross-national study. Educ. Psychol. Meas. 66, 701–716.

Tabachnick, B. G., & Fidell, L. S. (2007). Using multivariate statistics (5th ed.). Allyn & Bacon/Pearson Education.

Tatha, O., Shimazu, A., Watanabe, K., Kawakami, N., & Schaufeli, W. B. (2024). Measuring work engagement in Thailand: Development and validation testing of the Utrecht Work Engagement Scale-Thai version (UWES-TH). Industrial Health, 62(3), 182-194. https://doi.org/10.2486/indhealth.2023-0017

Willmer, M., Westerberg Jacobson, J., & Lindberg, M. (2019). Exploratory and Confirmatory Factor Analysis of the 9-Item Utrecht Work Engagement Scale in a Multi-Occupational Female Sample: A Cross-Sectional Study. Frontiers in Psychology, 10, 2771. https://doi.org/10.3389/fpsyg.2019.02771

Yang, P. J. (2017). Investigation and research on the current situation of learning burnout among students majoring in preschool education in secondary vocational schools [Master's thesis, Northwest Normal University].