ORIGINAL ARTICLE

Reliability and validity of the Vietnamese version of the 9-item Utrecht Work Engagement Scale

Thuy Thi Thu Tran MSc[1] | Kazuhiro Watanabe PhD[2] | Kotaro Imamura PhD[2] |
Huong Thanh Nguyen PhD[3] | Natsu Sasaki MD[2] | Kazuto Kuribayashi MHSc[4] |
Asuka Sakuraya PhD[5] | Nga Thi Nguyen MPH[3] | Thu Minh Bui MNSc[6] |
Quynh Thuy Nguyen PhD[1] | Tien Quang Truong PhD[3] | Giang Thi Huong Nguyen MNSc[6] |
Harry Minas RANZCP[7] | Akizumi Tsustumi MD[8] | Akihito Shimazu PhD[9] |
Norito Kawakami MD[2]

[1]Department of Occupational Health and Safety, Faculty of Environmental and Occupational Health, Hanoi University of Public Health, Hanoi, Vietnam

[2]Department of Mental Health, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan

[3]Faculty of Social Sciences – Behavior and Health Education, Hanoi University of Public Health, Hanoi, Vietnam

[4]Department of Psychiatric Nursing, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan

[5]Department of Public Health, Tokyo Women's Medical University, Tokyo, Japan

[6]Nursing Office, Bach Mai Hospital, Hanoi, Vietnam

[7]Melbourne School of Population and Global Health, The University of Melbourne, Melbourne, Vic., Australia

[8]Department of Public Health, Kitasato University School of Medicine, Sagamihara, Japan

[9]Faculty of Policy Management, Keio University, Fujisawa, Japan

Correspondence
Norito Kawakami, Department of Mental Health, Graduate School of Medicine, The University of Tokyo, 7-3-1, Hongo, Bunkyo-ku, Tokyo, 113-0033, Japan.
Email: nkawakami@m.u-tokyo.ac.jp

Abstract

Objectives: The present study investigated the reliability and validity of a newly developed Vietnamese version of the 9-item Utrecht Work Engagement Scale (UWES-9-V) in a sample of hospital nurses in Hanoi, Vietnam.

Methods: The UWES-9 was translated into Vietnamese following a standard procedure. A survey was conducted of 949 registered nurses in a large tertiary general hospital in Hanoi, Vietnam, in 2018, using a self-administered questionnaire including the Vietnamese UWES-9, other scales measuring health status, work performance, job demand, job control, and workplace social support, and questions pertaining to demographic variables. Cronbach's alpha and interclass correlation coefficients (ICC) were calculated to assess reliability. Explanatory and confirmatory factor analyses were conducted to assess factorial validity. Convergent validity was tested based on associations between the UWES-9-V and subscales and other scales.

Results: The Cronbach's alpha coefficients of the UWES-9-V and the Vigor, Absorption, and Dedication subscales were 0.93, 0.86, 0.77, and 0.90, respectively. ICC of the UWES-9-V in a subsample after 3 months was 0.48. Confirmatory factor analyses indicated an acceptable fit of both one-factor and three-factor structures, with the three-factor model having the better fit. The UWES-9-V and its subscales correlated with depression, anxiety and stress, health-related quality of life and health condition, job performance, and psychosocial work environment.

Conclusions: The study findings suggest that the UWES-9-V is a reliable and valid instrument to measure work engagement among hospital nurses in Vietnam, a low- and middle-income country. Future studies should confirm the validity and reliability of the UWES-9-V among various occupations.

1 | BACKGROUND

Since the introduction of positive psychology into occupational health, work engagement (WE), i.e., a positive, fulfilling, work-related state of mind that is characterized by vigor, dedication, and absorption, has received much attention, as it appears to be a key factor in the promotion of health and well-being and work performance among workers.[1] In the healthcare sector, nurses are an important workforce whose performance determines the quality of medical service provided. Higher levels of WE among nurses are associated with better patient outcomes,[2] and negatively related to burnout and psychological distress among nurses.[3,4] Even in low- and middle-income countries (LMICs), a shortage of healthcare professionals,[5] and increased demands for care from the aging population, nurses face increasing stress and burnout at work. This is particularly the case in South-East Asia[5] including Vietnam.[6] Research is needed to investigate the determinants and consequences of WE as well as an intervention to promote WE among nurses in LMICs in South-East Asia. However, only a few studies were reported on WE among nurses in these countries.[7,8] To promote research on WE among nurses in these LMICs, it is essential to establish a reliable and valid measure of WE among nurses in these countries.

Worldwide, several methods to measure WE are available in the literature.[9] Among them, the Utrecht Work Engagement scale (UWES)[10] is the most widely used tool to measure WE among nurses[9] as well as the general working population.[11] Between the original 17-item UWES (UWES-17) and the short version of the 9-item UWES (UWES-9), the latter scale is psychometrically better[11] and more stable in measuring WE over time.[12] The UWES-9 is widely used in studies of nurses in hospital settings since it requires less time and effort to complete[9] and has good reliability.[3,13] However, most validation studies of the UWES-9 among nurses (or multiple occupations including nurses) have been conducted in high-income countries[10,12-14] or in Europe.[12,13,15] No validation study of the UWES-9 among nurses in the South-East Asia region is available in the literature. Among LMICs in the Central and South Asia, the validity of the UWES-9 was only assessed in two countries, Nepal[3] and China.[4] In Vietnam, only one study used the UWES-9 to study WE among hospital nurses.[8] However, the translation did not follow a standard procedure, and the reliability and validity of that Vietnamese

UWES were not reported. The validity of the UWES-9 may be different across countries and cultures. For instance, the construct validity of the UWES-9 is inconsistent across published studies among nurses, with some reporting a one-factor model[14] but others reporting a three-factor model.[3,4] A formal Vietnamese version of the UWES-9 (UWES-9-V) needs to be developed and its validity tested among nurses in Vietnam in order to apply the concept of WE among nurses in Vietnam, and to compare the validity of the UWES-9-V with that of versions of other countries.

This study aimed to examine the reliability and validity of the UWES-9-V among hospital nurses in Vietnam. Reliability was assessed via internal consistency (Cronbach's alpha coefficient) and test-retest reliability (Intraclass correlation coefficient-ICC). Factorial validity and construct validity of the UWES-9-V were explored with confirmation factor analysis (CFA), explanatory factor analysis (EFA) and correlations with other scales.

2 | METHODS

2.1 | Participants and procedure

This study used data from a three-armed randomized controlled trial study that recruited full-time registered nurses in one large tertiary general hospital in Hanoi, Vietnam. Sampling and criteria to recruit participants were described by Imamura et al[16] A total of 1,258 eligible nurses were invited to participate in the study. A total of 949 of them returned the structured self-administrated questionnaires (75.44% response rate). Participants were equally and randomly assigned to two intervention groups and one control group based on their depression score. The original intervention study included a baseline survey (August-September 2018), a 3-month follow-up (January 2019), and a 7-month follow-up (April 2019).

A recent Risk of Bias checklist based on consensus-based standards for the selection of health measurement instruments (COSMIN) requires the minimum sample size for factor analysis to be 100.[17] In addition, the literature review shows that WE has a low correlation with other constructs (IrI ranged from 0.11 to 0.41),[3,18] and the minimum sample size is 853 when effect size = 0.1, alpha = 0.05, and 1-beta = 0.90, using G*Power version 3.1.9.4 (minimum sample size for Pearson's *r*, Exact test,

Correlation: Bivariate normal model. A priori: Computed required sample size, given alpha, power, and effect size). Therefore, data from the baseline survey (949 nurses) had an adequate sample size for the calculation of the UWES-9-V's Cronbach's alpha, CFA, EFA, correlations with other scales, and distribution of WE scores.

Data from baseline and 3-month follow-up surveys of a subsample (control group, 286 participants) were used to examine test-retest reliability with a 3-month interval.

2.2 | Measures

2.2.1 | The Vietnamese version of the 9-item Utrecht Work Engagement Scale(UWES-9-V)

An official Vietnamese version of the UWES-9 was not available when we started this study. We independently translated the original English version of the UWES-9[10] into Vietnamese, with the permission of Prof. Wilmar Schaufeli. The UWES-9 has a 7-point rating scale (0 = never; 6 = everyday) and consists of three questions to measure each of WE's three dimensions, namely, Vigor (items 1, 2, 5), Dedication (items 3, 4, 7), and Absorption (items 6, 8, 9). Subscale scores consist of the mean for each set of three questions. Hence, the three subscale scores range between 0 and 6. Similarly, the overall scale score consists of the mean of the nine items.

The development of the UWES-9-V followed the UWES's original manual[10] and the standard procedure of translation and cultural adaptation of patient reported health outcomes.[19] First, the forward translation of the scale from English to Vietnamese (draft UWES-9-V) was completed by one translator proficient in both languages with knowledge about the topic (April 2018). Face validity of the draft UWES-9-V was assessed among 30 nurses in the target hospital to test their understanding of the language translation. It also was sent to a psychiatrist for a professional assessment of the wording. Second, the draft UWES-9-V was piloted among 150 nurses (June, 2018). All comments and suggestions were carefully considered in the revision of the UWES-9-V. Third, the revised UWES-9-V was translated into English by a different proficient translator who did not participate in this project. The backward translation of the revised UWES-9-V was sent to its original author, Prof. Schaufeli, to compare with the original scale, followed by iterative processes. He raised questions on the translation of two items (item #2 and #9). Alternative wordings and phrases were proposed for these items based on an extensive discussion with the research team and senior nurses in the target hospital: the phrase of "strong and vigorous" (item #2) was replaced by a Vietnamese phrase meaning ""physically and mentally healthy, good at verbal communication,

and lively"; the phrase of "get carried away when working" (item #9) was replaced by a Vietnamese phrase meaning "I intensely concentrated at work so that I forgot other things". These amended items were again back-translated and approved by Prof. Schaufeli.

2.3 | Other scales for testing the construct validity of the UWES-9-V

2.3.1 | Depression, anxiety and stress

The 21-item version of the Depression, Anxiety and Stress Scale (DASS-21) measures symptoms of depression, anxiety, and stress in the community[20] with self-ratings on 21 statements such as "I was aware of dryness of my mouth." Each item is rated on a scale ranging from 0 (*did not apply to me at all*) to 3 (*applied to me very much, or most of the time*). The DASS-21 comprises three subscales with seven items for each dimension. The sum of the seven items' scores is the total score for each subscale, and thus ranges from 0 to 21.[20] The Vietnamese DASS-21 (DASS-21-V) was reported to be reliable in measuring mental problems among hospital nurses in Vietnam.[6] In this study, the Cronbach's alpha coefficients for the DASS-21-V and the Depression, Anxiety, and Stress subscales were 0.92, 0.82, 0.77, and 0.83, respectively.

2.3.2 | Health-related quality of life

The 5-level EQ-5D (EQ-5D-5L) is a widely used measurement of Health-related quality of life (HRQOL). The EQ-5D-5L contains five items covering mobility, self-care, usual activities, pain/discomfort, and anxiety/ depression, each of which is rated from level 1 (no problems) to level 5 (unable to). The Vietnamese version was developed and tested for its reliability and validity.[21] We used the Standard Value Set for Vietnam to calculate HRQOL scores for participants.[22]

2.3.3 | Psychosocial work environment

The Job Content Questionnaire (JCQ) is used to measure the psychosocial work environment.[23] The JCQ includes four subscales: a psychological demand scale (five items), a decision latitude (job control) scale (nine items), a supervisor support scale (four items), and a coworker support scale (four items), with a 4-point response options ranging from 1 (*strongly disagree*) to 4 (*strongly agree*). The reliability and validity of the Vietnamese version of JCQ among nurses in this study were acceptable.[24]

2.3.4 | Self-rated health

Participants were asked to rate their health status on a 100-point visual analog scale, in which 0 indicated the worst health and 100 meant the best health condition. This scale is a part of the EQ-5D-5L,[21] but not used in the computation of the HRQOL score.

2.3.5 | Global job performance

Global job performance was measured with the single item "*Using a 0-to-10 scale, how would you rate your overall job performance on the days you worked during the past 4 weeks (28 days)?*" It is a part of the World Health Organization—Work Health Performance Questionnaire (WHO HPQ), in which 0 indicates the "*worst possible work performance*" a person could have on this job and 10 means the "*top work performance*" on this job.[25] The final global job performance rating is the multiplication of this item score by 10, so this scale's score ranges from 0 to 100. The Vietnamese version of the WHO HPQ is not validated yet.

2.3.6 | Demographic variables

The following demographic variables were included in the analysis. Gender was categorized as male or female. Age in years was calculated based on the respondent's birth year. Three age groups were used, namely, under 30 years old (career start), 30-45 years old (transition to higher work position), and above 45 years old (prime working life and stable position at work). The labor contract was categorized into four types: fixed-term contract for less than 1 year, fixed-term contract for more than 1 year, no fixed-term contract, and permanent contract. The number of years working as a nurse was divided into three groups ($\leq$10, 11-20, and $\geq$21 years).

2.4 | Statistical analysis

Internal consistency reliability of the UWES-9-V and three subscales was evaluated using Cronbach's alpha coefficients. The mean and standard deviation (SD) of item scores, item-total correlation, and Cronbach's' alpha coefficient when an item was deleted for the UWES-9-V and each subscale were calculated. A Cronbach's alpha coefficient larger than 0.70 was acceptable.

Test-retest reliability of the UWES-9-V and subscales was examined using intraclass correlation coefficients (ICC), a two-way random model, and type absolute agreement. The 95% confidence interval (CI) of ICC values less than 0.50,

between 0.50 and 0.75, between 0.75 and 0.90, and greater than 0.90 were indicative of poor, moderate, good, and excellent reliability, respectively.[26] A paired *t*-test was conducted to compare the mean scores for the UWES-V and its subscales between two surveys.

For factorial validity, EFA was first conducted with maximum likelihood extraction. The number of factors was determined based on examination of the scree plot. Then, CFA (maximum likelihood estimation) was applied to examine the fit of one-factor and three-factor structure models. Several fit indices were used to determine the model fit, namely, the comparative fit index (CFI), the non-normed fit index (NNFI, also known as Tucker-Lewis index—TLV) greater than 0.95, the root-mean-square error of approximation (RMSEA) and 90% CI less than 0.06, standardized root-mean-square residual (SRMR) less than 0.08, the lower Akaike information criterion (AIC); chi-squared test, degrees of freedom (*df*), and significance level (*P*) were included.[27] Modification indices were examined if both structures did not have a good fit due to error covariance, and adjustments were made as necessary by freeing error covariance as reported by the previous studies.[12]

For convergent validity, we selected the following hypotheses to assess the associations of the UWES-9-V and subscales and other relevant scales. Based on the Job Demands-Resources (JD-R) model,[28] job resources (e.g., job control and social support at work from supervisors and coworkers) enhance WE.[29] The JD-R model did not address the association between job demand and WE; however, studies reported a weak positive association between these two constructs.[18] Regarding health outcomes, UWES-9 scores inversely and weakly correlated with depression, anxiety,[30] and psychological distress,[3] but was moderately and positively associated with health-related quality of life[31] and overall health.[3] UWES-9 scores also moderately and positively correlated with job performance.[3] In this study, these associations were assessed using Pearson's correlation (*r*).

Distribution of the UWES-9-V and its subscales' scores were examined based on gender, age, type of labor contract, and work experience using t-tests (for two groups) or one-way analysis of variance (ANOVA) and post hoc tests with Bonferroni correction (for more than two groups) at the significant level of *P* less than .05.

All analyses were conducted using the IBM Statistical Package for Social Sciences (SPSS) version 20 except for the confirmatory factor analysis, which was conducted using IBM SPSS analysis of moment structures (AMOS) version 20.

3 | RESULTS

Baseline data for 942 nurses were analyzed (final response rate 74.88%; 7 cases were removed because of missing data).

TABLE 1 Internal consistency reliability as Cronbach's alpha coefficients of UWES-9-V and subscales (N = 942)

Items	Mean	SD	Item-total correlation	Alpha coefficient when deleting an item
Item 1 bursting with energy (V1)	4.30	1.54	.72	.92
Item 2 strong and vigorous (V2)	4.47	1.44	.78	.92
Item 5 in the morning, feel like going to work (V3)	4.04	1.64	.79	.92
Item 3 enthusiastic about job (D1)	4.48	1.47	.85	.92
Item 4 job inspires me (D2)	4.25	1.48	.84	.92
Item 7 proud of the work (D3)	4.75	1.49	.77	.92
Item 8 immersed in work (A1)	4.31	1.46	.72	.92
Item 6 happy when working intensely (A2)	4.32	1.55	.82	.92
Item 9 get carried away when working (A3)	3.64	1.75	.45	.94
Vigor (UWES-V)	4.09	1.32	.86	
Dedication (UWES-D)	4.49	1.35	.90	
Absorption (UWES-A)	4.27	1.37	.77	
UWES-9-V	4.29	1.23	.93	

Abbreviations: SD: standard deviation; UWES-9-V: Vietnamese version of the 9-item Utrecht Work Engagement Scale; UWES-A: absorption subscale; UWES-D: dedication subscale; UWES-V: vigor subscale of UWES.

84.82% of respondents were female nurses with a mean age of 34.08 (SD = 6.79) and an average number of working years as a nurse of 10.33 (SD = 6.97). In the subsample of 286 nurses used to examine test-retest reliability, 84.27% of respondents were female nurses with mean age 33.81 (SD = 6.52) and an average number of working years as a nurse of 9.95 (SD = 6.61).

3.1 | Scale reliability

Table 1 presented information on the internal consistency of the UWES-9-V and the three subscales. The means of all items were identical, excepting item 9 (UWESA3 "*I get carried away when working*"), which had a lower mean score. This item also had a lower item-total correlation (0.45) compared with those of the other eight items (0.72–0.85). The Cronbach's alpha coefficients for the UWES-9-V, Vigor, Absorption, and Dedication subscales were 0.93; 0.86, 0.77, and 0.90, respectively. Most items substantially contributed to the whole scale's Cronbach's alpha coefficient, excepting item 9. The UWES-9-V's Cronbach's alpha increased 0.01 if item 9 was deleted.

Table 2 showed that the Vigor mean score significantly increased between the baseline and 3-month follow-up surveys (P = .04). There was no significant difference in mean scores for the UWES-9-V and the dedication, absorption subscales (P > .05). Test-retest reliability or stability of the UWES-9-V and three subscales was measured after 3 months in a subsample of 286 nurses (control group of the original intervention study). ICCs were calculated using two-way random effect models with absolute agreement. Stability of the UWES-9-V (95% CI) was 0.48 (0.35–0.59). The lowest value belonged to the Vigor subscale (0.35, 95% CI = 0.19–0.49).

3.2 | Scale validity

3.2.1 | Factorial validity

With the inconclusive agreement on factor validity of the UWES-9 and limited application of the UWES in LMICs, we examined both one-factor and three-factor models to see which models were appropriate within the Vietnamese context. In EFA, initial eigenvalues (% of variance explained)

TABLE 2 Test-retest reliability as IntraClass correlation coefficients of UWES-9-V and subscales in the 3-month interval (N = 286)

	Baseline		3 months		Paired t-test		Intraclass correlation	95% CI	
Items	Mean	SD	Mean	SD	t	P		Lower	Upper
Vigor	4.27	1.33	4.47	1.25	−2.09	.04	.35[**]	0.19	0.49
Dedication	4.62	1.23	4.70	1.25	−0.97	.33	.51[**]	0.39	0.61
Absorption	4.14	1.20	4.24	1.27	−1.19	.24	.50[**]	0.37	0.60
UWES-9-V	4.34	1.11	4.47	1.14	−1.64	.10	.48[**]	0.35	0.59

Abbreviations: CI, confidence interval, *p*, significant level; SD: standard deviation; *t*: *t*-test value.

**P < .01

TABLE 3 Fit indices of one-factor and three-factor models of the UWES-9-V: Confirmation factor analysis (N = 942)

Model	NNFI/ TLV	CFI	SRMR	RMSEA	RMSEA 90% CI	AIC	CMIN	df	P
M1—1 factor	0.91	0.92	0.13	0.15	0.14-0.16	636.04	600.04	27	<.01
M2—3 factor	0.93	0.93	0.12	0.15	0.13-0.16	541.27	499.27	24	<.01
M3—1 factor	0.97	0.97	0.06	0.09	0.08-0.10	262.41	222.41	25	<.01
M4—3 factor	0.97	0.98	0.05	0.09	0.08- 0.10	227.53	181.53	22	<.01

Abbreviations: AIC, Akaike information criterion; CFI, Comparative fit index, CI, Confidence interval; CMIN, Chi-square mean; df, degree of freedom; NNFI/ TLV, Non-normed fit index/Tucker-Lewis index; *P*: significant level; RMSEA, Root mean square error of approximation; SRMR, Standardized Root mean square residual.

for factor 1 to 9 were 5.95 (66.08), 0.93 (10.35), 0.57 (6.34), 0.35 (3.88), 0.31 (3.46), 0.29 (3.31), 0.23 (2.57), 0.21 (2.33), and 0.15 (1.68), respectively, yielding one factor. Preliminary CFA with robust maximum likelihood estimation (Models 1 and 2 in Table 3) showed that both the one- and three-factor models had a very poor fit with the current data. All fit indices were out of the acceptable range for both the one-factor model, where 9 items loaded on the uniform WE, and the three-factor model in which each item loaded on its theoretical dimensions (Vigor, Absorption, or Dedication) (NNFI/ TLV/CFI less than 0.95, SRMR and RMSEA exceeded 0.08 and 0.06, respectively). The analysis of modification indices showed that both models had substantial error covariance between two pairs of items: UWESV1 ("*Bursting with energy*") and UWESV2 ("*Strong and vigorous*"): UWESA1 ("*immersed in work*") and UWESA3 ("*Get carried away when working*") (error covariance ranged from 0.40 to 0.69 in Model 1 and from 0.29 to 0.60 in Model 2). After adjustment was made for the error covariance, the modified one-factor model (Model 3) and three-factor model (Model 4) obtained good fit indices, although RMSEA was slightly higher than the threshold 0.06 (Table 3).

Between the two modified models, the three-factor model showed a significantly better fit with the data (higher CFI, lower SRMR, smaller AIC, $\Delta\chi^2 = 100.87$, $\Delta df = 3$, $P < .01$). All factor loadings exceeded 0.70 in both modified models, excluding item 9 (UWESA3 "*Get carried away when working*") with loadings of 0.43 in Model 3 and 0.44 in Model 4. High correlations (>.90) between pairs of three latent

dimensions (vigor, absorption, and dedication) were presented in Figure 1.

3.2.2 | Convergent validity

Table 4 showed that UWES-9-V scores correlated negatively and weakly with scores for depression, anxiety, and stress ($r = −.33$ to $−.22$). It weakly and positively correlated with HRQOL, self-rated health, and work performance ($r = .26$, .28, and .21, respectively). Job demand weakly and negatively correlated with the UWES-9-V score ($r = −.16$), while the other three aspects of the psychosocial work environment, namely, job control, supervisor support, and coworker support positively correlated with WE ($r = .24$, .22, and .20, respectively).

3.3 | Distribution of WE among different groups

Table 5 presented comparisons of WE scores among demographic groups using t-tests or one-way ANOVAs with post hoc tests. A significant difference in WE was observed between males and females and among nurses with different labor contract types. Male nurses had considerably higher WE than their female coworkers, with their mean scores for the three dimensions of WE all being higher than those for the female nurses ($P < .05$). Nurses with labor contracts of
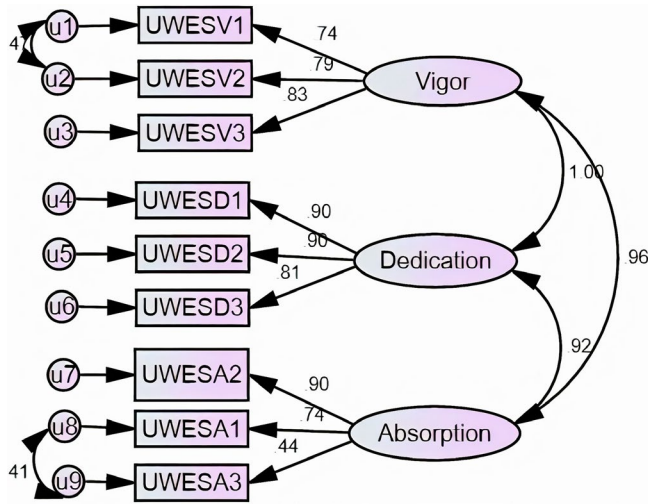
FIGURE 1 Modified path diagram of the UWES-9-V with standardized coefficients from the confirmatory factor analysis (N = 942). U: error, UWESV 1, 2, 3: items 1, 2, 3 of Vigor subscale or items 1, 2, 5 of the UWES-9-V; UWESD 1, 2, 3, of Dedication subscale or items 3, 4, 7 of the UWES-9-V; UWESA 1, 2, 3: items 1, 2, 3 of Absorption subscale or items 8, 6, 9 of the UWES-9-V.

less than 1 year expressed higher WE than those who signed no fixed-term contract with the hospital (P < .05).

4 | DISCUSSION

The Vietnamese version of the UWES-9 showed an acceptable level of reliability. The one- and three-factor structures of the UWES-9-V obtained acceptable fit, with the three-factor

model having a better fit. Convergent validity was supported with expected correlations between the UWES-9-V and other scales. The study findings suggest that the Vietnamese version of the UWES-9 is a reliable and valid instrument to measure WE among hospital nurses in Vietnam.

The new UWES-9-V had excellent internal consistency (Cronbach's alpha coefficient 0.93), higher than that was found in studies of nurses from Nepal (0.81),[3] one of the low-income countries. The UWES-9-V internal consistency was also better than that of versions of several other Asian countries, namely, Japan (0.92),[14] Malaysia (0.90),[7] and China (0.88).[4] Cronbach's alpha coefficients for the three subscales were acceptable but not as good as the uniform UWES-V-9. However, item 9 (UWESA3) showed considerably lower item-total correlation compared with the other items. This item also had low factor loading and noticeable error covariance in CFAs. Nevertheless, we kept item 9 since its presence had an insubstantial effect on scale reliability and validity. One possible reason for the differential performance of the item 9 could be that the Vietnamese wording that we selected for this item had a less positive nuance compared to that of other items and participants might rate this item differently from the other items. This item may measure a different construct, such as workaholism, rather than WE. Future studies should further examine the validity and reliability of item 9 in the UWES-9-V in different target groups. It is possible that both the contextually based translation and the direct meaning translation of this item (although a direct meaning translation might create confusion in the local language and context) should be examined in the validation process.[3]

Among previous studies, test-retest reliability of the UWES-9 varied, ranging from 0.66 (2-month interval)[14]

TABLE 4 Associations between UWES-9-V and three subscales and other scales

| Variables (score range) | Whole sample | | | Pearson's correlation coefficient[a] | | | |
	N[b]	Mean	SD	UWES-9-V	UWES-A	UWES-D	UWES-V
DASS depression (0-21)	927	3.00	2.90	−0.33	−0.20	−0.35	−0.35
DASS anxiety (0-21)	930	3.90	3.10	−0.22	−0.12	−0.26	−0.23
DASS stress (0-21)	929	5.60	3.60	−0.27	−0.16	−0.29	−0.28
HRQOL (0-1)	942	0.90	0.10	0.26	0.16	0.27	0.29
Self-rated health (0-100)	940	85.60	11.60	0.28	0.21	0.25	0.31
Global job performance (0-100)	942	84.40	11.20	0.21	0.19	0.18	0.20
JCQ job demand (12-48)	942	31.50	4.40	−0.16	−0.11	−0.18	−0.14
JCQ job control (24-96)	927	81.10	6.40	0.24	0.20	0.22	0.23
JCQ supervisor support (4-16)	937	12.00	1.90	0.22	0.17	0.21	0.22
JCQ coworker support (4-16)	940	12.20	1.50	0.20	0.16	0.19	0.21

Abbreviations: DASS, Depression, Anxiety and Stress scale; JCQ, Job Content Questionnaires; SD, standard deviation; UWES-9-V: Vietnamese version of the 9-item Utrecht Work Engagement Scale; UWES-V: vigor subscale of UWES; UWES-D, dedication subscale; UWES-A, absorption subscale.

[a]All correlations were significant at .01 level (one tailed).

[b]The number of participants varied due to missing responses on each variable.

TABLE 5 Work engagement across demographic groups (N = 942)

Variables	Group	N	UWES-9-V		UWES-A		UWES-D		UWES-V	
			Mean	SD	Mean	SD	Mean	SD	Mean	SD
Gender	Male	143	4.56[**]	1.16	4.35[*]	1.27	4.81[**]	1.26	4.51[*]	1.31
	Female	799	4.24	1.24	4.04	1.32	4.44	1.36	4.23	1.37
Types of labor contract	≤1 year	221	4.49[a,*]	1.07	4.27[a,*]	1.13	4.73[a,*]	1.18	4.46	1.22
	>1 year	29	4.26	1.15	4.06	1.16	4.39	1.44	4.33	1.18
	No fixed term	193	4.16	1.27	3.91	1.40	4.36	1.34	4.21	1.33
	Permanent	499	4.24	1.28	4.08	1.36	4.45	1.41	4.20	1.44
Age (years)	≤30	420	4.32	1.20	4.07	1.29	4.54	1.29	4.36	1.30
	31-45	479	4.24	1.27	4.10	1.35	4.44	1.41	4.18	1.42
	≥46	43	4.45	1.12	4.22	1.22	4.68	1.26	4.46	1.26
Years of working	≤10	545	4.34	1.17	4.10	1.28	4.56	1.26	4.34	1.29
	11-20	300	4.20	1.30	4.06	1.37	4.39	1.43	4.15	1.47
	≥21	97	4.27	1.33	4.11	1.39	4.44	1.52	4.27	1.45

Abbreviations: SD, standard deviation; UWES-9-V, Vietnamese version of the 9-item Utrecht Work Engagement Scale; UWES-V, vigor subscale of UWES; UWES-D, dedication subscale; UWES-A, absorption subscale.

[a]ANOVA with a post hoc test with Bonferroni correction. The mean scores of UWES, UWES-A, and UWES-D were significantly higher among participants with a labor contract less than 1 year as to compare with those among nurses with no fixed-term labor contract (P < .05). Otherwise, no significant difference (P > .05).

*P < .05.

**P < .01

to 0.70 (1-year interval)[10] and 0.91 (2-week interval).[32] The 3-month test-retest reliability (ICCs) in our study was moderate,[26] but lower than that of these reports. It was particularly lower for the Vigor subscale score, which significantly improved during the 3-month follow-up (from 4.27 to 4.47, P = .04). The results might be attributable to potential contamination of information on stress management in the present sample (the control group) from intervention groups who received smartphone-based cognitive-behavioral training programs.[16] Other possible explanations included a seasonal change in workload or other organizational changes in the target hospital. The stability and repeatability of the UWES-9-V needs more confirmation from future studies, with control of factors associated with WE.

The findings for the EFA and CFA indicated that both the one-factor and three-factor structures of the UWES-9-V were acceptable (CFI/ NNFI > 0.95, SRMR < 0.08, RMSEA close to 0.06). The three-factor model had a better fit for the current data but very high correlations between three dimensions (r > .90) indicated that the UWES-9-V could be used as a unitary construct to measure WE or applied separately to differentiate Vigor, Absorption, and Dedication. The factorial validity of the UWES-9-V was similar to the Nepalese UWES-9, which was validated among nurses in a LMIC country.[3] Most available studies supporting either the one-factor or three-factor structures were conducted among nurses in countries that were not LMICs.[4,10,12,14,15] Our findings contribute to the popularity and suitability of the UWES-9 with

the three-factor structure across diverse settings and occupations, including nurses in LMICs and Vietnam. However, the one-factor UWES-9 requires more robust evidence on its validity and reliability in measuring WE among the healthcare workforce in LMICs.

In our study, the UWES-9-V showed good convergent validity. Scores for depression, anxiety, and stress all correlated weakly and negatively with the UWES-9-V and its subscales' scores, which was concordant with previous findings of negative correlations between mental health problems and WE.[3,4,30] In previous studies, HRQOL and self-rated health also correlated moderately with WE.[3,31] Job control, supervisor support, and coworker support correlated weakly (r = .20-.24) in expected directions.[7,18] The negative correlation between job demand and WE in this study was inconsistent with findings from previous studies among Japanese employees[18] which reported low but positive correlations between WE and workload and time pressure. This could be attributable to cultural and occupational differences. High job demand is not regarded as motivation to work by Vietnamese nurses who have a high risk of increased psychological stress because of high workload and work pressure.[6] In addition, the target participants in this study were nurses in one hospital, while the Japanese studies were conducted with a considerably broader sample of the working population. Regarding performance, WE positively related to work performance among Vietnamese nurses. The work commitment of employees was associated with their performance and improvement of the

service they provided.[3] The present study confirms an adequate level of convergent validity of the UWES-9-V and subscales in measuring WE among hospital nurses.

The mean score of the UWES-9-V was lower than those in other studies among Malaysian nurses,[7] US emergency nurses,[33] and Finnish nurses.[34] It was considerably higher than results for Japanese nurses/employees[18,35] and Chinese registered nurses.[4] Lower WE scores in Japan have been attributed to a culture-specific response tendency to suppress positive emotions.[36] This may be not the case in Vietnam. When exploring WE among groups with different demographic characteristics, WE significantly associated with gender and type of labor contract. The considerably higher WE in men compared with women in our study was in line with that of other studies.[10] Female nurses reported lower scores than their male coworkers in all three WE dimensions as well as overall WE. In Vietnam, women are expected to be responsible for household chores and family care[37] while men hold the leading role in supporting the family. It is normal for men to rest and relax after work when their wives continue cooking, cleaning, or caring for other family members. In order to fulfill their traditional family responsibilities and also do their job, female nurses bear more workload and have comparably less leisure time in comparison with their male coworkers. This comparative lack of leisure time eventually affects both their physical and mental health, resulting in lower WE.[3,4] Hence, with a majority of employed women in this study (84.82%), support for female nurses at work is necessary to keep them engaged in hospital tasks.

Interestingly, nurses with short-term contracts were more enthusiastic about their job than those with no fixed-term contract. In Vietnam, having a contract without fixed-term brings more advantages for employees such as resignation with advance notice at any time for any reason without compensation related to breach of contract, and eligibility for unemployment insurance.[38] A contract without a fixed-term is usually used to replace short-term contracts for workers who have been working at the current workplace for over 3 years, but are still not eligible for a permanent contract by the time of contract renewal. This result is considered consistent with the decline in WE among nurses with more years of work experience in this study. Our results contradict prior research which reported a positive association between WE and work experience, and more commitment among permanent/ senior nurses in Nepal and Iran.[3,32] Our findings might be attributable to the typical work environment of Vietnamese tertiary hospitals. A cluster analysis reported that in comparison with younger nurses, older nurses with over 6 years of service experienced a combination of greater job demand, lower social support at work, and a higher risk of multiple mental disorders[6] which adversely affected their commitment to work. Our study emphasizes the importance of addressing the determinants of WE in occupational research to accurately and comprehensively interpret the study results.

The chief limitation of this study is the uniformity of participants, i.e., only nurses in one hospital, with an unequal distribution of gender. Caution should be exercised in generalizing these findings to the general population or other occupations in Vietnam. Second, the self-report approach may result in different understandings among study participants. Future studies should validate the UWES-9-V with a broader range of occupations. Third, one of the scales used to test the construct validity of the UWES-9-V, the work performance scale, has not been validated yet. Finally, we could not investigate other important properties of UWES-9-V such as content validity, cross-cultural validity, criterion validity, and responsiveness. However, the study clarifies the psychometric properties of the newly developed Vietnamese version of the UWES-9, a useful tool to measure WE among hospital nurses in Vietnam.

5 | CONCLUSION

This study confirms that the Vietnamese version of the UWES-9 has satisfactory psychometric properties and will be a suitable tool to measure WE in Vietnam, especially among nurses. This is also the first step to introducing the concept of positive occupational psychology, i.e., WE, in Vietnam in order to improve workers' mental health.

AUTHORS' CONTRIBUTIONS
KI, TTran, HN, KK, AS, TB, AN, QN, NN, GN, TTruong, HM, NS, AT, and NK conceived and designed the survey. TTran, HN, TB, AN, QN, NN, GN, and TTruong collected data. TTran and NK analyzed data and wrote the article. HN, AT, and AShimazu were major contributors to the study. KW, NS, AS, and NN provided important comments for the final manuscript. All authors read and approved the final manuscript.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE
The study protocol was reviewed and approved by the Research Ethics Committee of Graduate School of Medicine/Faculty of Medicine, The University of Tokyo

(no. 11991-(1)), and the Ethical Review Board for Biomedical Research, Hanoi University of Public Health (no 346/2018/YTCC-HD3) before implementation. All participants signed the consent form before joining this study.

DECLARATION OF INTERESTS

DATA AVAILABILITY STATEMENT

The datasets used for this study are available from the corresponding author on reasonable request.

ORCID

Thuy Thi Thu Tran https://orcid.org/0000-0002-2758-7695
Kazuhiro Watanabe https://orcid.org/0000-0002-3342-6142
Huong Thanh Nguyen https://orcid.org/0000-0002-9493-9590
Harry Minas https://orcid.org/0000-0003-1719-7367
Akizumi Tsustumi https://orcid.org/0000-0003-0966-4869
Akihito Shimazu https://orcid.org/0000-0002-7172-0043
Norito Kawakami https://orcid.org/0000-0003-1080-2720

REFERENCES

1. Schaufeli WB. The future of occupational health psychology. _Appl Psychol_. 2004;53(4):502–517.
2. Loerbroks A, Glaser J, Vu-Eickmann P, Angerer P. Physician burnout, work engagement and the quality of patient care. _Occup Med (Lond)_. 2017;67(5):356–362.
3. Panthee B, Shimazu A, Kawakami N. Validation of Nepalese version of Utrecht Work Engagement Scale. _J Occup Health_. 2014;56(6):421–429.
4. Fong TC, Ng SM. Measuring engagement at work: validation of the Chinese version of the Utrecht Work Engagement Scale. _Int J Behav Med_. 2012;19(3):391–397.
5. Elison NK, Verani AR, McCarthy C. National nursing and midwifery legislation in countries of South-East Asia with high HIV burdens. _WHO South-East Asia Journal of Public Health_. 2015;4(102):12–19.
6. Tran TTT, Nguyen NB, Luong MA, et al. Stress, anxiety and depression in clinical nurses in Vietnam: a cross-sectional survey and cluster analysis. _International Journal of Mental Health Systems_. 2019;13:3.
7. Othman N, Nasurdin AM. Social support and work engagement: a study of Malaysian nurses. _J Nurs Manag_. 2013;21(8):1083–1090.
8. Duong VMAT, Chitpakdee B Job Resources and Work Engagement among Nurses in Tertiary Care Hospitals, Thanh Hoa Province, The Socialist Republic of Vietnam. _Nursing Journal_. 2019;46(1), 14.
9. García-Sierra R, Fernández-Castro J, Martínez-Zaragoza F. Work engagement in nursing: an integrative review of the literature. _J Nurs Manag_ 2016;24(2):E101–E111.
10. Schaufeli WB, Bakker A. _Utrecht Work Engagement Scale: Preliminary Manual_. Utrecht, The Netherlands: Utrecht University; 2004. https://www.wilmarschaufeli.nl/publications/Schaufeli/Test%20Manuals/Test_manual_UWES_English.pdf.
11. Kulikowski K. Do we all agree on how to measure work engagement? Factorial validity of Utrecht Work Engagement Scale as a standard measurement tool – A literature review. _Int J Occup Med Environ Health_. 2017;30:161–175.
12. Seppälä P, Mauno S, Feldt T, et al. The construct validity of the Utrecht work engagement scale: multisample and longitudinal evidence. _J Happiness Stud_. 2008;10(4):459.
13. Sinval J, Marques-Pinto A, Queiros C, Maroco J. Work engagement among rescue workers: Psychometric properties of the Portuguese UWES. _Front Psychol_. 2017;8:2229.
14. Shimazu A, Schaufeli WB, Kosugi S, et al. Work engagement in Japan: Validation of the Japanese version of the Utrecht Work Engagement Scale. _Appl Psychol-Int Rev_. 2008;57(3):510–523.
15. Petrovic IB, Vukelic M, Cizmic S. Work Engagement in Serbia: Psychometric Properties of the Serbian Version of the Utrecht Work Engagement Scale (UWES). _Front Psychol_. 2017;8:1799.
16. Imamura K, Tran TTT, Nguyen HT, et al. Effects of two types of smartphone-based stress management programmes on depressive and anxiety symptoms among hospital nurses in Vietnam: a protocol for three-arm randomised controlled trial. _BMJ Open_. 2019;9(4):e025138.
17. Mokkink LB, de Vet HCW, Prinsen CAC, et al. COSMIN risk of bias checklist for systematic reviews of patient-reported outcome measures. _Qual Life Res_. 2018;27(5):1171–1179.
18. Inoue A, Kawakami N, Tsutsumi A, et al. Association of job demands with work engagement of Japanese employees: comparison of challenges with hindrances (J-HOPE). _PLoS One_. 2014;9(3):e91583.
19. Wild D, Grove A, Martin M, et al. Principles of good practice for the translation and cultural adaptation process for patient-reported outcomes (PRO) measures: Report of the ISPOR Task Force for Translation and Cultural Adaptation. _Value in Health_. 2005;8(2):94–104.
20. Lovibond PF, Lovibond SH. The structure of negative emotional states: comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. _Behav Res Ther_. 1995;33(3):335–343.
21. Tran BX, Ohinmaa A, Nguyen LT, Nguyen TA, Nguyen TH. Determinants of health-related quality of life in adults living with HIV in Vietnam. _AIDS care_. 2011;23(10):1236–1245.
22. Mai VQ, Sun S, Van Minh H, Luo N, Giang KB, Lindholm L, Sahlen KG, et al. An EQ-5D-5L Value Set for Vietnam. _Quality of Life Research_. 2020;29(7):1923–1933. http://dx.doi.org/10.1007/s11136-020-02469-7.
23. Karasek R, Brisson C, Kawakami N, Houtman I, Bongers P, Amick B. The Job Content Questionnaire (JCQ): an instrument for internationally comparative assessments of psychosocial job characteristics. _J Occup Health Psychol_. 1998;3(4):322–355.
24. Sasaki N, Imamura K, Thuy TTT, et al. Validation of the job content questionnaire among hospital nurses in Vietnam. _J Occup Health_. 2019;62(1):e12086.
25. Kessler RC, Barber C, Beck A, et al. The World Health Organization Health and Work Performance Questionnaire (HPQ). _J Occup Environ Med_. 2003;45(2):156–174.
26. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. _J Chiropr Med_. 2016;15(2):155–163.

27. Brown TA. *Confirmatory factor analysis for applied research*. New York, NY: Guilford Press; 2006.

28. Demerouti Evangelia, Bakker Arnold B., Nachreiner Friedhelm, Schaufeli Wilmar B. The job demands-resources model of burnout. *Journal of Applied Psychology*. 2001;86(3):499–512. http://dx.doi.org/10.1037/0021-9010.86.3.499.

29. Schaufeli WB, Bakker AB. Job demands, job resources, and their relationship with burnout and engagement: a multi-sample study. *J Organ Behav*. 2004;25(3):293–315.

30. Asghar H. Chapter 25: Patterns of engagement and anxiety in university students: first year to senior year. Students' work engagement & anxiety: Are they related? In: Pracana C ed. *Psychology Applications & Developments*. Lisboa, Portugal: inScience Press; 2014:248–259.

31. Mache S, Vitzthum K, Klapp BF, Danzer G. Surgeons' work engagement: Influencing factors and relations to job and life satisfaction. *The Surgeon*. 2014;12(4):181–190.

32. Torabinia M, Mahmoudi S, Dolatshahi M, Abyaz MR. Measuring engagement in nurses: the psychometric properties of the Persian version of Utrecht Work Engagement Scale. *Med J Islam Repub Iran*. 2017;31:15.

33. Siller J, Dolansky MA, Clavelle JT, Fitzpatrick JJ. Shared governance and work engagement in emergency nurses. *J Emerg Nurs*. 2016;42(4):325–330.

34. Mauno S, Ruokolainen M, Kinnunen U, De Bloom J. Emotional labour and work engagement among nurses: examining perceived compassion, leadership and work ethic as stress buffers. *J Adv Nurs*. 2016;72(5):1169–1181.

35. Kunie K, Kawakami N, Shimazu A, Yonekura Y, Miyamoto Y. The relationship between work engagement and psychological distress of hospital nurses and the perceived communication behaviors of their nurse managers: A cross-sectional survey. *Int J Nurs Stud*. 2017;71:115–124.

36. Shimazu A, Schaufeli WB, Miyanaka D, Iwata N. Why Japanese workers show low work engagement: An item response theory analysis of the Utrecht Work Engagement scale. *Biopsychosoc Med*. 2010;4:17.

37. Gender Equality Department - Ministry of Labour IaSA, United Nations Entity for Gender Equality and the Empowerment of Women. Facts and figures on women and men in Vietnam 2010-2015. Hanoi 2016.

38. Assembly VN. Labour code. In: Assembly VN ed. *Invalids and Social Affairs*. 10/2012/QH13. Vietnam: International Cooperation Department, Ministry of Labour; 2012.