



## ORIGINAL ARTICLE

# Psychometric properties of the Japanese version of the Occupational Fatigue Exhaustion Recovery Scale among shift-work nurses

Shinya Yamaguchi<sup>1</sup>  | Miho Sato<sup>2</sup> | Naomi Sumi<sup>2</sup> | Yoichi M. Ito<sup>3</sup> | Peter C. Winwood<sup>4</sup> | Rika Yano<sup>2</sup> 

<sup>1</sup>Graduate School of Health Sciences, Hokkaido University, Hokkaido, Japan

<sup>2</sup>Faculty of Health Sciences, Hokkaido University, Hokkaido, Japan

<sup>3</sup>Data Science Center, Promotion Unit, Institute of Health Science Innovation for Medical Care, Hokkaido University Hospital, Hokkaido, Japan

<sup>4</sup>School of Psychology, Social Work, and Social Policy, University of South Australia, Adelaide, South Australia, Australia

## Correspondence

Rika Yano, Faculty of Health Sciences, Hokkaido University, Hokkaido, Japan.  
Email: [r-yano@med.hokudai.ac.jp](mailto:r-yano@med.hokudai.ac.jp)

## Abstract

**Objectives:** Shift-work nurses are at a higher risk of inadequate recovery from fatigue and developing maladaptive fatigue with significant health consequences. Therefore, it is necessary to monitor fatigue and recovery levels with a reliable scale. We investigated psychometric properties of the Japanese version of the Occupational Fatigue Exhaustion Recovery scale (OFER-J) for shift-work nurses.

**Methods:** Japanese shift-work nurses responded to self-administered questionnaires at baseline ( $n = 942$ ) and one month later ( $n = 334$ ). The confirmatory factor analysis (CFA) and exploratory factor analysis (EFA) were conducted to verify the structural validity and the correlation analysis and one-way analysis of variance were conducted to test the construct and discriminative validity. Cronbach's alpha coefficient, intra-class correlation coefficient (ICC), and smallest detectable change (SDC) were calculated to assess reliability.

**Results:** The CFA showed high correlations between the factors and whilst the goodness-of-fit of the three-factor model was suboptimal, it was in an acceptable range. Most modifications included the error covariance of the Acute Fatigue (AF) and Intershift Recovery (IR) items. The EFA showed that Chronic Fatigue (CF) and AF were not clearly separated, indicating that the two AF items dropped out. Construct and discriminative validity were also well indicated. Cronbach's alpha coefficients were 0.75–0.85. Only CF showed sufficient reproductivity (ICC = 0.74). The SDC for CF, AF, and IR was 14.0, 17.1, and 18.7, respectively.

**Conclusions:** The validity and reliability of the OFER-J were verified as acceptable for shift-work nurses. The OFER-J could contribute to a data-based approach to fatigue management in nursing management practice.

## KEYWORDS

fatigue, Japan, psychometric, scale, shift-work nurse

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2022 The Authors. *Journal of Occupational Health* published by John Wiley & Sons Australia, Ltd on behalf of The Japan Society for Occupational Health

## 1 | INTRODUCTION

Shift-work is an essential work system in nursing to provide patients with 24-h care. However, this work system is a potential source of serious work-related fatigue. Previous studies have shown that nurses who work in rotation shift have higher levels of fatigue compared with day or fixed-shift nurses.<sup>1</sup> The impacts of nurses' fatigue are wide-ranging and serious problem: for example, previous studies have shown that nurses' fatigue reduces their work performance,<sup>2</sup> increases the risk of medical error,<sup>3</sup> and occupational injuries.<sup>4</sup> Maintaining safe working conditions within nursing requires appropriate management to minimize fatigue among shift-work nurses.

It is thus important to have scales that can reliably monitor fatigue within work settings. Work-related fatigue in nurses has been identified as both acute and chronic.<sup>5</sup> Depleted energy at the end of a work is evidence of acute fatigue. Comparatively, chronic fatigue manifests as a more serious and maladaptive fatigue state which can seriously undermine health. Chronic fatigue is a prolonged condition that causes negative emotions and inefficient action patterns in workers<sup>6</sup> and has more significant negative effects on nurses. Nursing tasks are physically and mentally demanding, and shift-work nurses are required to work long shifts (e.g., 12 h), resulting in high levels of acute fatigue.<sup>2</sup> Acute fatigue is generally relieved by regular rest.<sup>7</sup> However, shift-work in nursing often involves short periods of rest between shifts (quick return) due to overtime and irregular shift schedules.<sup>8</sup> Such working conditions cause poor sleep quality<sup>9</sup> compromising recovery from fatigue. Due to acute fatigue and low levels of recovery, shift-work nurses are at a higher risk of developing chronic fatigue.<sup>10</sup> Among nurses, quick return increases the risk of accidents that harm nurses, patients, and facility equipment<sup>11</sup> due to fatigue and sleep deprivation. Hence, to minimize these effects, the monitoring of interrelated acute fatigue, chronic fatigue, and recovery levels is needed to detect the signs of the accumulation of fatigue at an earlier stage.

In previous studies, the Multidimensional Fatigue Inventory,<sup>12</sup> the Chalder Fatigue Scale,<sup>13</sup> the Fatigue Assessment Scale,<sup>14</sup> the Jikaku-sho shirabe,<sup>15</sup> and the Cumulative Fatigue Symptom Index<sup>16</sup> have been used to assess nurses' fatigue. The Need for Recovery Scale<sup>17</sup> and the Recovery Experience Questionnaire<sup>18</sup> have been used to measure recovery. However, these scales were developed to measure either acute fatigue, chronic fatigue, or recovery levels. Combining them in one study leads to a burdensome long questionnaire. Therefore, a brief scale capable of measuring all three factors simultaneously was needed.

Winwood et al. (2005) developed the Occupational Fatigue Exhaustion Recovery scale (OFER) to measure acute fatigue, chronic fatigue, and intershift recovery

levels.<sup>19</sup> A revised version of the OFER was published in 2006.<sup>6</sup> The OFER is a brief scale that imposes a lesser burden on respondents and can be easily adapted to work settings. To date, translations of the OFER have been developed for nurses in Lebanon,<sup>20</sup> China,<sup>21</sup> and Korea,<sup>22</sup> and the psychometric properties of these versions have been validated. However, the Japanese version of OFER (OFER-J) has not yet been developed. In Japan, the length of shifts has transitioned to longer hours, with many nurses working 16-h night shifts, which are unprecedented in other countries. Even more challenging, 21.0% of nurses reported that they had to undertake consecutive shifts with quick return in a month.<sup>23</sup> A further survey revealed that 41.5% of nurses reported experiencing rest intervals of <8 h between shifts.<sup>24</sup> Such work patterns seriously compromise nurses' health and patients' safety. It is, therefore, imperative to develop the OFER-J, and to inform strategies to reduce fatigue among Japanese shift-work nurses.

### 1.1 | Objectives

We developed the OFER-J and examined its validity and reliability for registered nurses and midwives engaged in shift-work in hospitals.

### 1.2 | Hypotheses

We hypothesized that the OFER-J would have adequate internal consistency, test-retest reliability, and a three-factor model structure. According to Winwood et al., each subscale of the OFER measures the following<sup>6,19</sup>:

1. Chronic Fatigue (CF): a complex of mental, physical, and emotional component (including depressive element).
2. Acute Fatigue (AF): an inability and/or unwillingness to engage with normal activities (including self-chosen pleasure activities) as a direct consequence of previous activity.
3. Intershift Recovery (IR): the extent to which acute work-related fatigue is perceived to have been recovered, or dissipated, by the time the next work shift commences.

To test construct validity, the Japanese version of the Multidimensional Fatigue Inventory (MFI-J),<sup>25</sup> the Short Form Health survey-36 (SF-36-J),<sup>26</sup> and the Pittsburg Sleep Quality Index (PSQI-J)<sup>27</sup> were used. For the CF, we considered that the MFI as a multidimensional chronic fatigue scale<sup>12</sup> and the Mental Health included in the SF-36 would be useful.<sup>6,19</sup> Similarly, the

Vitality and other subscales included in the SF-36-J would be useful for the AF. Regarding the IR, related studies<sup>21,22</sup> reported correlations between the IR and the global score of the PSQI. Thus, the following hypotheses are proposed:

1. The CF score is positively correlated with the MFI-J total score and negatively correlated with the SF-36-J Mental Health score.
2. The AF score correlates negatively with the following subscales of the SF-36-J: Vitality, Role Physical, Role Emotional, and Social Functioning.
3. The IR scores are negatively correlated with the global score of the PSQI-J.

We assumed that all correlations will be moderate or high (correlation coefficient  $\geq 0.30$ ).<sup>28</sup> Additionally, based on previous studies,<sup>2,29,30</sup> the discriminative validity of the OFER-J was tested by comparing work-related characteristics among the participants.

## 2 | MATERIAL AND METHODS

### 2.1 | Development of Japanese version of the OFER

The permission to develop the OFER-J was obtained from the original author (PCW).<sup>6</sup> Four researchers (SY, MS, NS, and RY) in nursing translated the OFER into Japanese. SY is a doctoral student with a master's degree in nursing and a registered nurse. MS is a lecturer with a Ph.D. in health sciences, NS is an associate professor with a Ph.D. in health sciences, and RY is a professor with a Ph.D. in nursing. Each researcher independently translated the OFER, and the results were compared and discussed to ensure that the text and content were appropriate and understandable to nurses. When all four researchers agreed, a provisional OFER-J was produced. Next, a back translation of the provisional OFER-J was undertaken by an English native translator, while simultaneously checking for cultural equivalence. The OFER-J was sent to the original author to confirm the semantic equivalence several times until their final approval was obtained. The contents of each item were reviewed and validated by five researchers in nursing with clinical experience.

### 2.2 | Study design and participants

This study using an anonymous self-administered questionnaire was conducted from January to March 2021.

A total of 1,188 nurses working in fifty-five wards in six hospitals from three cities in Japan were included in the study. To examine the test-retest reliability, the participants in one out of six hospitals were asked to complete the questionnaire twice. They were asked to complete the second OFER-J one month after the first distribution. Two things were considered when the interval of one month was decided: the OFER is a scale that measures the state of fatigue and recovery in the two preceding months<sup>6</sup> and the effect of the respondents' memory. The two responses were corresponded using anonymous IDs comprising continuous numbers assigned to the questionnaires.

The participants were registered nurses and midwives who provide direct patient care, were full-time employees, and engaged in shift-work in the inpatient wards. The exclusion criteria were as follows: (1) nurse managers; (2) outpatient, laboratory, and operating departments; (3) nurses with less than one year of experience; and (4) nurses on leave. The sample size was calculated to be  $n = 910$  for the confirmatory factor analysis to test the structural validity, with  $\alpha = 0.05$ , power = 0.80, root-mean-square error of approximation = 0.040, and degrees of freedom = 87.<sup>31</sup> The response rate was assumed as 80%, and the final number of distributions was set at 1250, considering the missing data.

This study was approved by the Ethical Review Board of the University to which the authors are affiliated (approval number: 20-49) and was conducted in accordance with the Declaration of Helsinki.

### 2.3 | Data collection

The questionnaires were distributed to all participants by the administrators. The location and timing of the responses to the questionnaire were not specified. After completing the questionnaire, the participants sealed the questionnaire in an envelope and deposited it in a box placed in each ward. The response period was set for two weeks after distribution.

### 2.4 | Instruments

#### 2.4.1 | Occupational Fatigue Exhaustion Recovery Scale

We used the OFER-J as described above. The scale consists of 15 items, with 5 items each for CF, AF, and IR. All items are based on a 7-point Likert-type scale (0 = *strongly disagree* to 6 = *strongly agree*), and items 9, 10, 11, 13, and 15 were reverse-coded. The following formula was

used to calculate the standardized score (range, 0–100) for each subscale from the response results:

$$\text{sum (each subscale's item scores) / 30} \times 100.$$

For AF and CF, a higher standardized score indicated a higher degree of chronic or acute fatigue, while a higher standardized score for IR indicated a better degree of recovery.

## 2.4.2 | Multidimensional Fatigue Inventory

We used the MFI-J which has been verified for reliability and validity among Japanese workers.<sup>25</sup> The MFI-J consists of 20 items with five subscales. Total scores (range, 20–100), where higher values indicate more severe fatigue, were used in this analysis. In this study, Cronbach's alpha coefficient was 0.88.

## 2.4.3 | Pittsburgh Sleep Quality Index

Subjective sleep quality was measured using the PSQI-J.<sup>27</sup> This study only used the 18 items that were answered by the participants. The PSQI-J, comprising of seven components (range of subscale scores = 0–3), and the global PSQI-J score (range: 0–21), which is the sum of these components, was used in the analysis. The higher the global PSQI-J score, the poorer the sleep quality. In this study, Cronbach's alpha coefficient was 0.67.

## 2.4.4 | Short-form health survey 36

The SF-36-J, consisting of 36 items, measures health-related QOL in eight domains.<sup>26</sup> We used the following five domains: Role Physical, Vitality, Social Functioning, Role Emotional, and Mental Health. The scores for each domain were expressed on a scale from 0 to 100, with lower scores indicating poorer conditions. A web-based scoring program was used to calculate the scores for each domain, and norm-based scores were used. In this study, Cronbach's alpha coefficient was 0.77–0.93.

## 2.4.5 | Participant characteristics and work conditions

The sex, age, current profession, years of nursing experience, educational level, and family roles of participants were measured. Work conditions, such as ward, shift type, and number of night shifts (previous month), were assessed.

## 2.5 | Statistical analysis

Statistical analysis was performed using JMP Pro software, ver. 15.0 (SAS Institute Inc., Cary, NC, USA). The significance level was set at 5%. In terms of descriptive statistics, the mean and standard deviation (*SD*) and frequencies (percentages) were used.

Cronbach's alpha coefficients were calculated to assess the internal consistency of each subscale with a cut-off of 0.70.<sup>32</sup> In each subscale of the OFER-J, if those who fell into the minimum (0 point) and maximum value (100 points) exceeded 15%, it was considered a floor and ceiling effect, respectively.

The confirmatory factor analysis (CFA) was conducted to verify the structural validity of the OFER-J. The cut-off for the factor loadings of the items was set at 0.30.<sup>33</sup> In evaluating the fit of the model, the  $\chi^2$  value was not employed as a goodness-of-fit indicator in this study because it is always significant when the sample size is large.<sup>34</sup> Instead, the following indices were used: goodness-of-fit index (GFI > 0.90), normed fit index (NFI > 0.90), comparative fit index (CFI > 0.90), Tucker-Lewis index (TLI > 0.90), root-mean-square error of approximation (RMSEA < 0.080), and standardized root-mean-square residual (SRMR < 0.080).<sup>34,35</sup> In cases where the CFA did not show a good model fit, an exploratory factor analysis (EFA) was performed. A parallel analysis<sup>36</sup> and Velicer's minimum average partial (MAP) test<sup>37</sup> were used to determine the factor structure more precisely. To obtain the factor structure, a principal component analysis (PCA) using the maximum likelihood method was performed and the Quartimin rotation method, an oblique rotation, was used for factor extraction.

To test the construct validity, correlation coefficients (*r*) between the OFER-J and the MFI-J Total score, the global PSQI-J score, and each subscale score of the SF-36-J were calculated using Spearman's rank-order correlation analysis with a listwise method. We also compared the OFER-J subscales among participants and work-related characteristics by using one-way analysis of variance (ANOVA). Tukey's honestly significant difference test was used for multiple comparisons.

To examine the test-retest reliability, the intraclass correlation coefficient (ICC) was calculated by a two-way random effect model for all subscales and items, with a cut-off of 0.70.<sup>32</sup> The standard error of measurement (SEM) was calculated as the square root of the sum of the between-measures variance and the residual variance. Smallest detectable change (SDC) was calculated based on the SEM, using the formula<sup>32</sup>

$$\text{SDC} = 1.96 \times \sqrt{2} \times \text{SEM}.$$

### 3 | RESULTS

#### 3.1 | Response, participants, and work-related characteristics

Data from 942 participants (valid response rate: 79.3%) were used for the analysis (Appendix Figure S1). Responses to the second test were obtained from 351 participants, of which 334 valid responses were used for the analysis.

The mean (*SD*) age of the participants was 33.8 (9.7) years. Among the participants, 87.3% were female, and 72.7% were in their 20s and 30s. Two-shift schedules (52.8%) were more common than three-shift schedules (47.2%) (Table 1).

#### 3.2 | Internal consistency and floor/ceiling effect

Cronbach's alpha coefficients for the subscales were as follows: CF = 0.75, AF = 0.85, and IR = 0.84. There was no floor and/or ceiling effect for any subscale.

#### 3.3 | Validity

##### 3.3.1 | Structural validity

The data were poorly fitted to the three-factor model: except for SRMR (=0.067), none of the fit indices showed acceptance criteria, GFI = 0.866, NFI = 0.846, CFI = 0.855, TLI = 0.825, RMSEA = 0.115 (95% confidence interval [CI]: 0.109–0.121). The factor loadings for all items were above 0.30. Therefore, the model was modified without removing any of the items.

Based on the modified indices of CFA, it was shown that adding covariance between the item errors, as shown in Figure 1, improved the model fit. A total of 12 error covariances were added, indicating that the data fit a three-factor structure in an acceptable manner.

However, as there was a lot of error covariance among the factors, we performed an EFA to reveal a detailed data structure. There were six factors for the parallel test and two for the MAP test. Considering scale reliability, we performed a PCA and Quartimin rotation to extract 2–4 factors (Table 2).

In the two-factor solution, Factor 1 included all IR items with factor loadings above 0.30; the others included two AF (AF4 and AF5) and three CF items (CF1, CF3, and CF4). Factor 2 contained AF and CF items, while IR3 showed cross loadings on Factor 2. Both factors contained a mixture of all subscales' items.

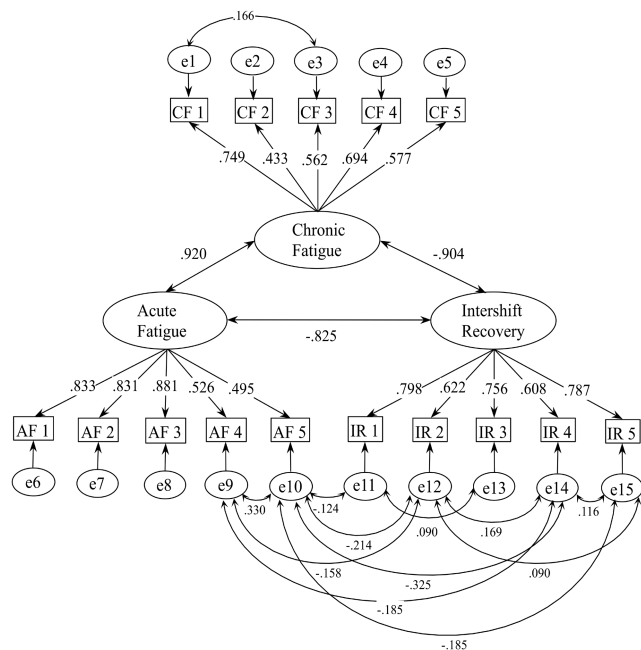
Considering the three-factor solution, Factor 1 also included three items from the AF (AF1–AF3) and three

TABLE 1 Participant and work-related characteristics (*n* = 942)

	Number (%)
Age (%) <sup>a</sup>	
20–29	426 (45.4)
30–39	256 (27.3)
40–49	168 (17.9)
50–59	83 (8.9)
≥60	5 (0.5)
Sex (%)	
Female	822 (87.3)
Male	120 (12.7)
Current profession (%)	
Registered nurse	884 (93.8)
Midwife	58 (6.2)
Educational level (%) <sup>b</sup>	
High school (Five-year)	26 (2.8)
Vocational school	504 (53.8)
Junior college	50 (5.3)
University	326 (34.8)
Graduate school	31 (3.3)
Marital status (%) <sup>c</sup>	
Married	279 (29.6)
Unmarried (single/divorced)	662 (70.4)
With pre-school child (%) <sup>d</sup>	
Yes	84 (9.0)
No	849 (91.0)
Care role at home (%) <sup>e</sup>	
Yes	25 (2.7)
No	907 (97.3)
Hospital	
University hospital	425 (45.1)
Municipal hospital	191 (20.3)
General hospital	326 (34.6)
Wards	
Medical/surgical/mix	597 (63.4)
Psychiatry	53 (5.6)
Maternity	58 (6.2)
High care unit/emergency	137 (14.5)
Palliative care	13 (1.4)
Others	84 (8.9)
Shift type	
8 h night duty (three-shift schedule)	445 (47.2)
12 h night duty (two-shift schedule)	245 (26.0)
16 h night duty (two-shift schedule)	252 (26.8)
Number of night shifts (last month, three-shift schedule) <sup>f</sup>	
1–2	25 (5.8)
3–5	277 (64.1)
6–8	127 (29.4)
≥9	3 (0.7)
Number of night shifts (last month, two-shift schedule) <sup>g</sup>	
1–2	30 (6.1)
3–4	176 (35.9)
≥5	285 (58.0)

Note: Others include comprehensive rehabilitation ward, wards for community-based care, and sanatorium long-term care ward.

<sup>a</sup>*n* = 938. <sup>b</sup>*n* = 937. <sup>c</sup>*n* = 941. <sup>d</sup>*n* = 933. <sup>e</sup>*n* = 932. <sup>f</sup>*n* = 432. <sup>g</sup>*n* = 491.



Model fit indices:

Goodness-of-fit index = 0.940

Normed fit index = 0.932

Comparative fit index = 0.940

Tucker-Lewis index = 0.916

Root-mean-square error of approximation = 0.079 (95% CI: 0.073–0.086)

Standardized root-mean-square residual = 0.046

**FIGURE 1** Results of confirmatory factor analysis with standardized estimates and error covariances for the Japanese version of the Occupational Fatigue Exhaustion Recovery scale ( $n = 942$ ). Abbreviations: AF, Acute Fatigue; CF, Chronic Fatigue; IR, Intershift Recovery

items from the IR (IR1, IR3, IR5) subscale, but the factor loadings for all CF items exceeded 0.30, indicating that this factor was related to CF. Similarly, Factor 2 was related to IR. However, these factors included items from different subscales. Factor 3 only included items from the AF (AF3–AF5) and IR (IR2 and IR4) subscales.

Even in the four-factor solution, there was a factor (Factor 1) that contained a mixture of AF and CF items. Similar to the three-factor solution, AF4 and AF5 dropped out of Factor 1 and were included in Factor 3.

### 3.3.2 | Construct and discriminative validity

Table 3 shows the correlation matrix between the OFER-J subscales and other scales. There was a strong positive correlation between the CF subscale and MFI-J Total ( $r = 0.653$ ), and a strong negative correlation between the Mental Health subscale of the SF-36-J ( $r = -0.573$ ). For the AF subscale, there was a moderate to strong negative correlation between the SF-36-J, especially with Vitality

( $r = -0.650$ ). As hypothesized, there was a moderate negative correlation between the IR subscale and the global PSQI-J score ( $r = -0.412$ ).

One-way ANOVA revealed statistically significant differences in OFER-J scores based on age, family role (marital status and with pre-school child), and shift type (Table 4).

## 3.4 | Reliability

### 3.4.1 | Test-retest reliability and measurement error

Table 5 shows the results of test-retest reliability, SEM, and SDC for all subscales and items. Only the CF subscale had an ICC above 0.70.

## 4 | DISCUSSION

In Japan, most shift-work nurses are in their 20s or 30s, approximately 90% are female,<sup>23</sup> and most shift types are two-shift schedule, which is consistent with our results. Therefore, the participants in this study were a sample with similar characteristics to shift-work nurses in Japanese hospitals.

The internal consistency of the OFER-J based on the scale structure of the original version was adequate. Thus, the items included in each factor were all homogeneous and consistently reflected the properties of AF, CF, and IR, respectively.<sup>38</sup> Therefore, the OFER-J can be score in a way similar to that in the original version.

In the CFA, although the model fit met the all goodness-of-fit indices through modification, the standardized estimates among factors were higher than in previous studies.<sup>6,21,22</sup> In the EFA (all solutions), factors with a mixture of AF and CF items were identified, indicating that the association between AF and CF was particularly strong for Japanese shift-work nurses. There are some possible reasons for this. Unlike in the original<sup>6</sup> and Chinese versions,<sup>21</sup> the participants in this study were all engaged in shift-work with night shifts. Additionally, in contrast to the Korean version,<sup>22</sup> more than half of our participants were engaged in a two-shift schedule with long working hours. Further, Japanese workers may be inherently prone to overworking, as expressed by the term Karoshi,<sup>39</sup> which was created by the Japanese work culture. Japanese shift-work nurses provide care for over 10 patients on long night shifts.<sup>30</sup> These unique work cultures may have the coexistence of AF and CF without mediated recovery due to the high workload and fatigue for Japanese shift-work nurses, making it difficult to clearly distinguish between

**TABLE 2** Factor loadings of the 2–4 factor model in the OFER-J (principal component analysis, maximum likelihood method, quartimin rotation method)

Item/factor	Two factor solution		Three factor solution			Four factor solution			
	Factor 1	Factor 2	Factor 1	Factor 2	Factor 3	Factor 1	Factor 2	Factor 3	Factor 4
Ofer-AF1		0.79	0.77			0.73			
Ofer-AF2		0.69	0.71	−0.31		0.60			
Ofer-AF3		0.73	0.75		0.33	0.63			
Ofer-AF4	−0.50				0.71			0.66	
Ofer-AF5	−0.80			−0.49	0.65			0.77	
Ofer-CF1	−0.45	0.40	0.49	−0.58		0.31			0.60
Ofer-CF2		0.59	0.51			0.52			
Ofer-CF3	−0.41		0.33	−0.52					0.60
Ofer-CF4	−0.31	0.45	0.50	−0.34		0.30	−0.34		
Ofer-CF5		0.53	0.52			0.35			
Ofer-IR1	0.55		−0.41	0.58			0.76		
Ofer-IR2	0.65			0.52	−0.34		0.35		
Ofer-IR3	0.38	−0.43	−0.51	0.42			0.85		
Ofer-IR4	0.80			0.58	−0.41			−0.35	−0.35
Ofer-IR5	0.64		−0.36	0.73			0.37		−0.47
Total cumulative variance (%)	57.2			63.3			68.9		

Note: Items with factor loadings below 0.30 are not shown.

Abbreviations: AF, Occupational Fatigue Exhaustion Recovery scale-Acute Fatigue; CF, Occupational Fatigue Exhaustion Recovery scale-Chronic Fatigue; IR, Occupational Fatigue Exhaustion Recovery scale-Intershift Recovery.

**TABLE 3** Correlation coefficients between OFER-J, MFI, PSQI, and SF-36

	Ofer-AF <sup>a</sup>	Ofer-CF <sup>a</sup>	Ofer-IR <sup>a</sup>	MFI total <sup>b</sup>	PSQI-G <sup>c</sup>	SF-36 RP <sup>d</sup>	SF-36 RE <sup>e</sup>	SF-36 SF <sup>a</sup>	SF-36 VT <sup>f</sup>	SF-36 MH <sup>f</sup>
Ofer-AF	1.000									
Ofer-CF	0.688***	1.000								
Ofer-IR	−0.730***	−0.637***	1.000							
MFI Total	0.551***	0.653***	−0.557***	1.000						
PSQI-G	0.356***	0.378***	−0.412***	0.455***	1.000					
SF-36 RP	−0.467***	−0.434***	0.391***	−0.499***	−0.321***	1.000				
SF-36 RE	−0.437***	−0.477***	0.408***	−0.518***	−0.381***	0.680***	1.000			
SF-36 SF	−0.353***	−0.349***	0.343***	−0.400***	−0.308***	0.458***	0.495***	1.000		
SF-36 VT	−0.650***	−0.562***	0.595***	−0.725***	−0.450***	0.473***	0.496***	0.434***	1.000	
SF-36 MH	−0.460***	−0.573***	0.465***	−0.598***	−0.416***	0.455***	0.580***	0.473***	0.633***	1.000

Note: Abbreviations: OFER-AF, Occupational Fatigue Exhaustion Recovery scale-Acute Fatigue; OFER-CF, Occupational Fatigue Exhaustion Recovery scale-Chronic Fatigue; OFER-IR, Occupational Fatigue Exhaustion Recovery scale-Intershift Recovery; MFI Total, Multidimensional Fatigue Inventory-Total score; PSQI-G, Global score of Pittsburgh Sleep Quality Index; SF-36, Short-Form Health Survey 36; RP, Role-Physical; RE, Role-Emotional; SF, Social Functioning; VT, Vitality; MH, Mental Health.

Correlation coefficients were calculated by Spearman's rank-order correlation analysis with listwise method that excludes from the calculation cases where any one of the multiple variables to be analyzed has a missing value.

<sup>a</sup>*n* = 942. <sup>b</sup>*n* = 930. <sup>c</sup>*n* = 862. <sup>d</sup>*n* = 940. <sup>e</sup>*n* = 939. <sup>f</sup>*n* = 940. \*\*\**p* < 0.001.

the two constructs. In the future, we might confirm a factor structure that clearly distinguishes between AF and CF by including workers who have lower workloads.

IR items were consistent even when the number of factors was increased in the EFA; in the three-factor solution, the factor loadings of the five IR items were

TABLE 4 Comparison of OFER-J subscale scores by demographic and work-related groups

	OFER-CF		OFER-AF		OFER-IR	
	mean (SD)	<i>p</i>	mean (SD)	<i>p</i>	mean (SD)	<i>p</i>
Age						
20–29 ( <i>n</i> = 426)	53.4 (19.5)	0.138	63.7 (18.2)	0.749	36.0 (18.9)	0.013
30–39 ( <i>n</i> = 256)	51.0 (19.5)		63.0 (18.9)		34.7 (19.6)	
40–49 ( <i>n</i> = 168)	51.0 (16.6)		64.6 (17.4)		37.8 (19.2)	
≥ 50 ( <i>n</i> = 88)	49.0 (19.4)		62.3 (20.0)		42.2 (21.9)	
Marital status						
Married ( <i>n</i> = 279)	48.8 (18.6)	0.001	62.3 (18.2)	0.177	37.9 (19.9)	0.163
Unmarried ( <i>n</i> = 662)	53.2 (19.1)		64.1 (18.5)		35.9 (19.3)	
With pre-school child						
Yes ( <i>n</i> = 84)	45.7 (19.0)	0.002	61.2 (17.8)	0.220	40.2 (20.8)	0.069
No ( <i>n</i> = 849)	52.4 (18.9)		63.8 (18.4)		36.2 (19.3)	
Care role at home						
Yes ( <i>n</i> = 25)	57.2 (14.4)	0.146	68.0 (18.9)	0.217	34.0 (16.9)	0.500
No ( <i>n</i> = 907)	51.6 (19.1)		63.4 (18.4)		36.7 (19.5)	
Shift type						
8 h night duty <sup>a</sup> ( <i>n</i> = 445)	51.9 (19.4)	0.003	63.3 (18.0)	0.031	34.7 (19.1)	0.002
12 h night duty <sup>b</sup> ( <i>n</i> = 245)	54.9 (18.2)		66.0 (18.0)		36.3 (19.4)	
16 h night duty <sup>b</sup> ( <i>n</i> = 252)	49.0 (18.8)		61.7 (19.4)		40.0 (19.8)	
Number of night shifts <sup>c</sup>						
1–2 ( <i>n</i> = 25)	57.2 (18.5)	0.178	66.1 (13.3)	0.536	36.4 (18.0)	0.697
3–5 ( <i>n</i> = 277)	52.3 (19.4)		63.7 (19.0)		34.9 (19.7)	
≥ 6 ( <i>n</i> = 130)	49.8 (62.2)		62.2 (16.9)		33.5 (18.2)	
Number of night shifts <sup>d</sup>						
1–2 ( <i>n</i> = 30)	49.2 (21.0)	0.629	65.6 (18.7)	0.454	36.9 (21.4)	0.712
3–4 ( <i>n</i> = 176)	51.7 (18.9)		62.5 (19.3)		39.0 (19.8)	
≥ 5 ( <i>n</i> = 285)	52.5 (18.3)		64.6 (18.4)		37.6 (19.3)	

Note: OFER-AF, Occupational Fatigue Exhaustion Recovery scale-Acute Fatigue; OFER-CF, Occupational Fatigue Exhaustion Recovery scale-Chronic Fatigue; OFER-IR, Occupational Fatigue Exhaustion Recovery scale-Intershift Recovery; SD, standard deviation.

*P*-values for between-group comparisons were assessed by one-way analysis of variance.

<sup>a</sup>*p* < 0.050.

<sup>b</sup>*p* < 0.010 (Tukey's honestly significant difference test).

<sup>c</sup>Three-shift schedule.

<sup>d</sup>Two-shift schedule.

<sup>e</sup>Three-shift schedule (last month).

<sup>f</sup>Two-shift schedule (last month).



**TABLE 5** Test-retest reliability, standard of measurement error, and smallest detectable change of the OFER-J

	Score range	Overall ( <i>n</i> = 942) mean (SD)	Test-retest survey ( <i>n</i> = 334)				
			Test 1 mean (SD)	Test 2 mean (SD)	ICC	SEM	SDC
OFER-CF	0–100	51.9 (19.0)	50.4 (19.4)	49.4 (19.4)	0.74	5.1	14.0
CF 1	0–6	3.4 (1.6)	3.2 (1.6)	3.3 (1.6)	0.70	0.9	2.4
CF 2	0–6	2.0 (1.8)	1.7 (1.7)	1.7 (1.6)	0.64	1.0	2.7
CF 3	0–6	3.9 (1.7)	3.9 (1.7)	3.8 (1.6)	0.69	0.9	2.6
CF 4	0–6	3.3 (1.6)	3.3 (1.6)	3.2 (1.6)	0.62	1.0	2.7
CF 5	0–6	3.0 (1.5)	3.0 (1.5)	2.9 (1.4)	0.60	0.9	2.6
OFER-AF	0–100	63.6 (18.4)	61.8 (19.0)	56.0 (12.3)	0.62	6.2	17.1
AF 1	0–6	3.5 (1.5)	3.3 (1.5)	3.2 (1.5)	0.65	0.9	2.5
AF 2	0–6	4.3 (1.3)	4.2 (1.4)	4.1 (1.3)	0.67	0.8	2.1
AF 3	0–6	3.6 (1.5)	3.5 (1.5)	3.4 (1.5)	0.63	0.9	2.5
AF 4	0–6	3.6 (1.4)	3.5 (1.3)	2.6 (1.3)	0.46	1.8	5.0
AF 5	0–6	4.1 (1.4)	4.1 (1.2)	3.4 (1.3)	0.45	1.0	2.8
OFER-IR	0–100	36.5 (19.5)	37.4 (19.5)	42.7 (12.3)	0.54	6.8	18.7
IR 1	0–6	2.1 (1.6)	2.1 (1.5)	2.0 (1.3)	0.40	1.1	3.0
IR 2	0–6	2.2 (1.4)	2.4 (1.4)	3.7 (1.5)	0.29	1.9	5.4
IR 3	0–6	2.1 (1.6)	2.3 (1.6)	2.4 (1.6)	0.67	0.9	2.5
IR 4	0–6	2.3 (1.5)	2.3 (1.4)	2.4 (1.4)	0.57	0.9	2.5
IR 5	0–6	2.2 (1.5)	2.2 (1.5)	2.3 (1.4)	0.59	0.9	2.6

Note: OFER-AF, Occupational Fatigue Exhaustion Recovery scale-Acute Fatigue; OFER-CF, Occupational Fatigue Exhaustion Recovery scale-Chronic Fatigue; OFER-IR, Occupational Fatigue Exhaustion Recovery scale-Intershift Recovery; SD, standard deviation; ICC, intra-class correlation coefficient, SEM, standard error of measurement; SDC, smallest detectable change.

sufficient, despite cross loadings. Therefore, Factor 2 was related to IR.

The results of the CFA showed that there was an error covariance between AF4 and AF5, consistent with the results of the original<sup>6</sup> and Chinese versions.<sup>21</sup> Additionally, these items required error covariances with multiple items in the IR items. AF4 (“I usually have lots of energy to give to my family or friends”) and AF5 (“I usually have plenty of energy left for my hobbies and other activities after I finish work”) contained phrases indicating how to spend time outside of work. The term “spending time with family or friends” and “engaging in my hobbies and other activities” have been reported to be effective in coping with and recovering from stress and fatigue in Japan and elsewhere.<sup>29,40</sup> From the above things, AF4 and AF5 were related to IR for Japanese nurses. Further, the EFA (three-factor solution) revealed that AF4 and AF5 were not included in the factors of other AF items. However, because of the limitations of many factors such as the condition of the participants, additional validation should be conducted to determine whether two items should be excluded from the AF.

Our results suggest that the association between the three factors (especially CF and AF) may be strong for Japanese

shift-work nurses; therefore, careful interpretation must be made to clearly distinguish each factor from the scoring results. In addition, two AF items dropped out of the factors that included other AF items, but the decision to exclude these items should be based on further studies considering the measurement reliability. Furthermore, since the factor analysis showed cross loadings of multiple items, it may be useful to apply item scores along with subscale scores.

For construct validity, the associations between the OFER-J and other scales were in the direction and magnitude as hypothesized. These results indicate that the OFER-J subscales have the capacity to measure CF, AF, and IR, respectively. The results showed significant associations with several variables, supporting the discriminative validity of the OFER-J. First, the association of the participants and work-related characteristics was consistent with the previous study.<sup>29,30</sup> Second, the three subscale scores significantly differed by shift type. The AF scores of those on the 16-h night shift (two-shift schedule) were significantly lower than of those on the 12-h night shift (two-shift schedule). In Japan, a nap break of coherent duration is recommended for the 16-h night shift,<sup>23</sup> and it is possible that this nap influenced this result. Additionally, this shift type was characterized by the ease of taking a coherent rest

day.<sup>23</sup> Thus, nurses engaged in this shift had the highest IR score and a significantly lower CF score.

The results of test-retest reliability showed that only CF had an ICC were above 0.70, indicating sufficient test-retest reliability. The ICCs for AF and IR were below 0.70, consistent with the results of previous studies,<sup>21,22</sup> which did not show sufficient test-retest reliability. According to Winwood et al. (2005), AF and IR reflect acute or subchronic conditions that vary with work, week, and nonwork-time behaviors.<sup>19</sup> Therefore, it is reasonable to consider that AF and IR subscales measure temporally variable states rather than stable traits. The SDC is a criterion to determine whether the observed change is a real change beyond the measurement error.<sup>32</sup> If the change exceeds the SDC, it indicates that a “real change” has occurred in the individual. When evaluating the effects of interventions, it is recommended to observe changes in the OFER-J score and to use SDC as a reference value.

There are limitations to this study. The main issue is sample representativeness. The characteristics of the participants were similar to those of shift-work nurses in Japanese hospitals; however, since this study was conducted in one region, the results might not be generalized. Further studies need to be conducted for a more diverse pool of participants. The validation with nurses working in other institutions (e.g., elderly care facilities) will also be necessary. Additionally, studies are needed to verify the applicability of the OFER-J to other healthcare professionals such as caregivers and physicians to increase its generalizability. Second, some psychometric properties, such as face validity, cross-cultural validity, and responsiveness, were not examined. Third, as all variables in our study were obtained by self-report measures, recall bias and desirability bias could not be ruled out. In addition, to ensure power, all correlation analyses were performed using a listwise method, but the selection bias was not addressed. Finally, SARS-Cov-2 infection continued to spread when this study was conducted. However, it was difficult to assess which participants were responsible for caring for COVID-19 patients and the impact that had on working conditions; therefore, the possible impact of this on the results could not be addressed.

## 5 | CONCLUSION

In Japanese shift-work nurses, the OFER-J had good internal consistency and test-retest reliability. In addition, construct and discriminative validity were well indicated. However, care should be taken in interpreting the three factors independently. Considering the importance of understanding and addressing the fatigue and recovery level of shift-work nurses, this study is the first step toward

supporting a data-driven approach to fatigue management. The OFER-J is useful regarding (1) surveillance and monitoring of fatigue and recovery levels and (2) assessing the effect of shift scheduling design. For Japanese shift-work nurses, administrators, and policymakers, the OFER-J could contribute to fatigue risk management.

## DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

## ACKNOWLEDGMENTS

The authors thank all the nurses and nurse managers for their participation in this study.

## CONFLICT OF INTEREST

There are no conflicts of interest to declare.



## AUTHOR CONTRIBUTIONS

S.Y., M.S., N.S., and R.Y. contributed to the conceptualization and design of this study; S.Y. and R.Y., contributed to the data collection; S.Y. contributed to the data curation and analysis; Y.M.I. provided advice with and reviewed the statistical analysis; S.Y. conducted the statistical analysis and drafted the manuscript; S.M., N.S., Y.M.I., P.C.W., and R.Y. provided academic suggestions and reviewed the manuscript; and RY was supervised throughout the whole study process and responsible for the final content. All authors read and approved the final manuscript.

## INFORMED CONSENT

An explanatory document was attached stating that participation in the study was voluntary and that no one would be disadvantaged if they refused to participate. Informed consent was obtained by submitting the completed questionnaire form.

## ORCID

Shinya Yamaguchi  <https://orcid.org/0000-0001-6802-2837>  
Rika Yano  <https://orcid.org/0000-0002-5243-8763>

## REFERENCES

1. Ferri P, Guadi M, Marcheselli L, Balduzzi S, Magnani D, Di Lorenzo R. The impact of shift work on the psychological and physical health of nurses in a general hospital: a comparison between rotating night shifts and day shifts. *Risk Manag Healthc Policy*. 2016;9:203-211. [10.2147/RMHP.S115326](https://doi.org/10.2147/RMHP.S115326)
2. Barker LM, Nussbaum MA. Fatigue, performance and the work environment: a survey of registered nurses. *J Adv Nurs*. 2011;67(6):1370-1382. [10.1111/j.1365-2648.2010.05597.x](https://doi.org/10.1111/j.1365-2648.2010.05597.x)

3. Abdulah DM, Perot KA, Holroyd E. Fatigue as a primary and secondary factor in relation to shift-rotating and patient safety in nurses. *J Hosp Adm.* 2019;9(1):1. [10.5430/jha.v9n1p1](https://doi.org/10.5430/jha.v9n1p1)
4. Smith DR, Mihashi M, Adachi Y, Nakashima Y, Ishitake T. Epidemiology of needlestick and sharps injuries among nurses in a Japanese teaching hospital. *J Hosp Infect.* 2006;64(1):44-49. [10.1016/j.jhin.2006.03.021](https://doi.org/10.1016/j.jhin.2006.03.021)
5. Smith-Miller CA, Shaw-Kokot J, Curro B, Jones CB. An integrative review: Fatigue among nurses in acute care settings. *J Nurs Adm.* 2014;44(9):487-494. [10.1097/NNNA.000000000000104](https://doi.org/10.1097/NNNA.000000000000104)
6. Winwood PC, Lushington K, Winefield AH. Further development and validation of the occupational fatigue exhaustion recovery (OFER) scale. *J Occup Environ Med.* 2006;48(4):381-389. [10.1097/01.jom.0000194164.14081.06](https://doi.org/10.1097/01.jom.0000194164.14081.06)
7. Chen J, Davis KG, Daraiseh NM, Pan W, Davis LS. Fatigue and recovery in 12-hour dayshift hospital nurses. *J Nurs Manag.* 2014;22(5):593-603. [10.1111/jonm.12062](https://doi.org/10.1111/jonm.12062)
8. Flo E, Pallesen S, Moen BE, Waage S, Bjorvatn B. Short rest periods between work shifts predict sleep and health problems in nurses at 1-year follow-up. *Occup Environ Med.* 2014;71(8):555-561. [10.1136/oemed-2013-102007](https://doi.org/10.1136/oemed-2013-102007)
9. Dahlgren A, Tucker P, Gustavsson P, Rudman A. Quick returns and night work as predictors of sleep quality, fatigue, work-family balance and satisfaction with work hours. *Chronobiol Int.* 2016;33(6):759-767. [10.3109/07420528.2016.1167725](https://doi.org/10.3109/07420528.2016.1167725)
10. Fang J, Qiu C, Xu H, You G. A model for predicting acute and chronic fatigue in Chinese nurses. *J Adv Nurs.* 2013;69(3):546-558. [10.1111/j.1365-2648.2012.06029.x](https://doi.org/10.1111/j.1365-2648.2012.06029.x)
11. Vedaa Ø, Harris A, Erevik EK, et al. Short rest between shifts (quick returns) and night work is associated with work-related accidents. *Int Arch Occup Environ Health.* 2019;92(6):829-835. [10.1007/s00420-019-01421-8](https://doi.org/10.1007/s00420-019-01421-8)
12. Smets EMA, Garssen B, Bonke B, De Haes JCJM. The multidimensional fatigue inventory (MFI) psychometric qualities of an instrument to assess fatigue. *J Psychosom Res.* 1995;39(3):315-325. [10.1016/0022-3999\(94\)00125-O](https://doi.org/10.1016/0022-3999(94)00125-O)
13. Chalder T, Berelowitz G, Pawlikowska T, et al. Development of a fatigue scale. *J Psychosom Res.* 1993;37(2):147-153. [10.1016/0022-3999\(93\)90081-P](https://doi.org/10.1016/0022-3999(93)90081-P)
14. Michielsen HJ, De Vries J, Van Heck GL. Psychometric qualities of a brief self-rated fatigue measure. *J Psychosom Res.* 2003;54(4):345-352. [10.1016/S0022-3999\(02\)00392-6](https://doi.org/10.1016/S0022-3999(02)00392-6)
15. Kubo T, Tachi N, Takeyama H, et al. Characteristic patterns of fatigue feelings on four simulated consecutive night shifts by “Jikaku-sho Shirabe”. *SANGYO EISEIGAKU ZASSHI.* 2008;50(5):133-144. [10.1539/sangyoeisei.B7008](https://doi.org/10.1539/sangyoeisei.B7008). (in Japanese).
16. Kosugo R. Validity and reliability of cumulative fatigue symptoms index. *J Sci Labour.* 1991;67(4):145-157. (in Japanese).
17. van Veldhoven M. Measurement quality and validity of the “need for recovery scale”. *Occup Environ Med.* 2003;60:3i-9. [10.1136/oem.60.suppl\\_1.i3](https://doi.org/10.1136/oem.60.suppl_1.i3)
18. Sonnentag S, Fritz C. The recovery experience questionnaire: Development and validation of a measure for assessing recuperation and unwinding from work. *J Occup Health Psychol.* 2007;12(3):204-221. [10.1037/1076-8998.12.3.204](https://doi.org/10.1037/1076-8998.12.3.204)
19. Winwood PC, Winefield AH, Dawson D, Lushington K. Development and validation of a scale to measure work-related fatigue and recovery: The Occupational Fatigue Exhaustion/Recovery Scale (OFER). *J Occup Environ Med.* 2005;47(6):594-606. [10.1097/01.jom.0000161740.71049.c4](https://doi.org/10.1097/01.jom.0000161740.71049.c4)
20. Clinton M, Tchapanian H. Rasch analysis of Lebanese nurses' responses to the Occupational Fatigue Exhaustion Recovery Scale. *J Nurs Meas.* 2017;25(3):459-475. [10.1891/1061-3749.25.3.459](https://doi.org/10.1891/1061-3749.25.3.459)
21. Fang JB, Zhou CF, Huang J, Qiu CJ. Psychometric properties of the Chinese version of the Occupational Fatigue Exhaustion/Recovery scale: A test in a nursing population. *J Nurs Res.* 2018;26(3):191-198. [10.1097/jnr.0000000000000247](https://doi.org/10.1097/jnr.0000000000000247)
22. Min A, Min H, Hong HC. Psychometric properties of the Korean version of the Occupational Fatigue Exhaustion Recovery Scale in a nurse population. *Res Nurs Health.* 2019;42(5):358-368. [10.1002/nur.21980](https://doi.org/10.1002/nur.21980)
23. Japanese Nursing Association. Survey of nursing in hospitals and medical clinics with bed 2019; [online] 2020. [cited 2021 Aug 15]; Available from <https://www.nurse.or.jp/home/publication/research/index.html> (in Japanese)
24. Japan Federation of Medical Worker's Unions. Night shift survey report 2020; [Online] 2020. [cited 2021 Aug 15]; Available from <http://irouren.or.jp/research/324e9714869db7acc3f8b80e6c3206459a902144.pdf> (in Japanese)
25. Sugaya N, Kaiya H, Iwasa R, Nomura S. Reliability and validity of the Japanese version of Multidimensional Fatigue Inventory (MFI). *Job Stress Res.* 2005;12:233-240. (in Japanese).
26. Fukuhara S, Bito S, Green J, Hsiao A, Kurokawa K. Translation, adaptation, and validation of the SF-36 Health Survey for use in Japan. *J Clin Epidemiol.* 1998;51(11):1037-1044. [10.1016/S0895-4356\(98\)00095-X](https://doi.org/10.1016/S0895-4356(98)00095-X)
27. Doi Y, Minowa M, Okawa M, Uchiyama M. Development of the Japanese version of the Pittsburgh Sleep Quality Index. *Jpn J Psychiatr Treat.* 1998;13(6):755-763 (in Japanese).
28. Cohen J. A power primer. *Psychol Bull.* 1992;112(1):155-159. [10.1037//0033-2909.112.1.155](https://doi.org/10.1037//0033-2909.112.1.155)
29. Winwood PC, Winefield AH, Lushington K. Work-related fatigue and recovery: The contribution of age, domestic responsibilities and shiftwork. *J Adv Nurs.* 2006;56(4):438-449. [10.1111/j.1365-2648.2006.04011.x](https://doi.org/10.1111/j.1365-2648.2006.04011.x)
30. Fang J, Kunaviktikul W, Olson K, Chontawan R, Kaewthummanukul T. Factors influencing fatigue in Chinese nurses. *Nurs Health Sci.* 2008;10(4):291-299. [10.1111/j.1442-2018.2008.00407.x](https://doi.org/10.1111/j.1442-2018.2008.00407.x)
31. Hancock GR, Freeman MJ. Power and sample size for the root mean square error of approximation test of not close fit in structural equation modeling. *Edu Psychol Meas.* 2001;61(5):741-758. [10.1177/00131640121971491](https://doi.org/10.1177/00131640121971491)
32. Terwee CB, Bot SDM, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol.* 2007;60(1):34-42. [10.1016/j.jclinepi.2006.03.012](https://doi.org/10.1016/j.jclinepi.2006.03.012)
33. Watson R, Thompson DR. Use of factor analysis in journal of advanced nursing: Literature review. *J Adv Nurs.* 2006;55(3):330-341. [10.1111/j.1365-2648.2006.03915.x](https://doi.org/10.1111/j.1365-2648.2006.03915.x)
34. Hooper D, Coughlan J, Mullen M. Structural equation modeling: Guidelines for determining model fit. *Electron J Bus Res Methods.* 2008;6:53-60. [10.21427/D7CF7R](https://doi.org/10.21427/D7CF7R)
35. McDonald RP, Ho MHR. Principles and practice in reporting structural equation analyses. *Psychol Methods.* 2002;7(1):64-82. [10.1037/1082-989X.7.1.64](https://doi.org/10.1037/1082-989X.7.1.64)

36. Horn JL. A rationale and test for the number of factors in factor analysis. *Psychometrika*. 1965;30(2):179-185. [10.1007/BF02289447](https://doi.org/10.1007/BF02289447)
37. Velicer WF, Eaton CA, Fava JL. *Construct explication through factor or component analysis: A review and evaluation of alternative procedures for determining the number of factors or components*. Goffin RD, Helmes E, eds. Springer; 2000.
38. Streiner DL, Norman GR. *Health measurement scales: A practical guide to their development and use*, 4th ed. Oxford University Press; 2008.
39. Takahashi M. Sociomedical problems of overwork-related deaths and disorders in Japan. *J Occup Health*. 2019;61(4):269-277. [10.1002/1348-9585.12016](https://doi.org/10.1002/1348-9585.12016)
40. Kubo T, Takahashi M, Sallinen M, Kubo Y, Suzumura H. How are leisure activity and shiftwork schedule associated with recovery from fatigue in shiftwork nurses? *SANGYO EISEIGAKU*

*ZASSHI*. 2013;55(3):90-102. [10.1539/sangyoeisei.B12004](https://doi.org/10.1539/sangyoeisei.B12004) (in Japanese).

## SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

**How to cite this article:** Yamaguchi S, Sato M, Sumi N, Ito YM, Winwood PC, Yano R. Psychometric properties of the Japanese version of the Occupational Fatigue Exhaustion Recovery Scale among shift-work nurses. *J Occup Health*. 2022;64:e12325. doi:[10.1002/1348-9585.12325](https://doi.org/10.1002/1348-9585.12325)