

Psychometric Properties of a New Simplified Scale to Assess the Quality of Workers' Rest and Recreation on Their Days Off

Osamu Itani,* Yoshitaka Kaneita,* Satoru Harano,* Mikiko Tokiya,† Yuichiro Otsuka,* Yuuki Matsumoto,* Sachi Nakagome* and Yu Kinoshita*

*Division of Public Health, Department of Social Medicine, Nihon University School of Medicine, Itabashi-ward, Tokyo 173-8610, Japan, and †Department of Public Health and Epidemiology, Faculty of Medicine, Oita University, Yufu-city, Oita 879-5593, Japan

ABSTRACT

Background In this study, we developed and tested the validity and reliability of the 12-item Rest and Recreation Quality Scale (RRQ-Scale) for Workers as a new scale capable of conveniently assessing the quality of workers' days off, that is, their rest and recreation.

Methods Participants included 756 employees (694 men, 62 women, mean age \pm SD = 44.7 ± 13.5 , age range = 18–81) of 26 manufacturing-related companies located in Oita Prefecture, Japan. We analyzed the factorial validity of the scale's score distribution and its criterion-related validity and reliability (Cronbach's coefficient α), compared to MOS 36-Item Short-Form Health Survey version 2.0 (SF-36v2) and the generalized self-efficacy scale. Participants responded to the 12 questions on 4-point scales; these were summed to calculate the total score (score range: 12–48 points).

Results The smallest and largest mean \pm SD of the questions were 2.53 ± 0.89 and 3.21 ± 0.74 , respectively, which are within the possible score range (1–4, implying the absence of floor and ceiling effects). A confirmatory factor analysis indicated that the goodness of fit of the higher-order factor analysis model was satisfactory (GFI = 0.955), confirming factorial validity. In addition, consistent with the theoretical predictions, the total score exhibited statistically significant positive correlations with the components of the SF-36v2; physical component summary (PCS): $r_s = 0.193$, $P < 0.001$, mental component summary (MCS): $r_s = 0.369$, $P < 0.001$, and role/social Component Summary (RCS): $r_s = 0.115$, $P = 0.002$. This confirmed criterion-related validity. Further, the overall reliability of the scale was high; $\alpha = 0.877$.

Conclusion In sum, the score distribution, validity, and reliability of the RRQ-Scale for workers were good, indicating a high degree of practicality.

Key words hobbies; leisure activities; quality of life; relaxation; work-life balance

Maintenance of worker health is a major theme in the field of occupational health. Health problems are associated with reduced work ability¹ and labor productivity,² as well as with risks, such as absenteeism due to

illness.³ Thus, they can be an important determinant of starting and continuing to work.⁴ In particular, “presenteeism” is defined as a state of reduced productivity due to poor physical condition or health circumstances.⁵ Recently, research on this concept has been identified as an important theme in occupational health.

Research into the factors associated with presenteeism is being pursued through a variety of perspectives. In addition to workers' social and work environments (such as high demands, heavy workloads, time pressure, shift-work, and long working hours^{6–8}), lifestyle (such as sleep duration, physical inactivity, poor diet, and smoking^{2,9}) has also been associated with presenteeism. When considering workers' lives overall, the polar opposites of work are “rest,” in which the mind and body recuperate from the fatigue of work, and “recreation,” which involves mental and physical activities that increase mental vitality and health. Rest and recreation include various elements, including sleep habits and how days off are spent. Previous research has focused on how work and sleep impact health and presenteeism,^{6,8,9} but few studies have investigated how workers spend their days off. One reason this construct has not been studied might owe to the lack of indices and scales capable of evaluating their quality. Therefore, we attempted to create an assessment scale based on a previous study¹⁰ capable of assessing it. Our assessment scale has the two factors of “rest” and “recreation” as important structural components of days off. “Rest” during days off refers to how the day off is spent, the purpose of which is to recover from the physical and mental fatigue by being free from labor or work. On

Corresponding author: Yoshitaka Kaneita, MD, PhD

nusmpublichealth@gmail.com

Received 2020 October 26

Accepted 2020 December 21

Online published 2021 January 18

Abbreviations: AGFI, adjusted goodness of fit index; CFI, comparative fit index; GFI, goodness of fit index; HRQOL, health-related quality of life; MCS, mental component summary; PCS, physical component summary; QOL, quality of life; RCS, role/social Component Summary; RMR, root mean square residual; RMSEA, root mean square error of approximation; SF-36v2, MOS 36-Item Short-Form Health Survey version 2.0

the other hand, “recreation” includes the practical ways days off are spent, the purpose of which is to create meaning in life by engaging in activities that provide spiritual nourishment, such as hobbies and social activities. A questionnaire developed in a previous study¹⁰ had the main constructs of “rest” and “recreation” in days off and used 16 questions to qualitatively evaluate these two elements. However, this previous study did not report its reliability or validity, nor the distribution of scores, and thus did not provide sufficient evidence for using it as a scale. In addition, 16 is a relatively large number of questions, which creates problems regarding the scale’s practicality. Therefore, in the present study, we decided to review the questions and carefully screen them to create a more practical questionnaire and collect new data to verify the scale’s reliability and validity (factorial and criterion-related validity) to establish evidence for using it. As a result, we confirmed that this newly created assessment scale was sufficiently practical and named it “the Rest and Recreation Quality Scale (RRQ-Scale) for Workers.” Japan is in the midst of a major shift in labor policy under the name of “Work System Reform.”¹¹ In these circumstances, such a scale could be an important tool.

MATERIALS AND METHODS

Basic attributes

We collected data on survey subjects’ sex (male or female) and age as basic attributes. We also collected data on worker type (blue-collar or white-collar), job category (management or nonmanagement), whether they performed shift work, amount of overtime per month, and number of days off per month.

Composition of scale

The RRQ-Scale focuses on two factors regarding how workers spend their time outside of work: physical and mental recovery (rest factor) and cultivating vitality through the enjoyment of hobbies, sports, and leisure (recreation factor). Thus, it assesses the quality of how workers spend their days off through a self-administered questionnaire. The candidate items for the RRQ-Scale were developed in a previous study.¹⁰ Candidate questions were considered by drafting questions pertaining to factors promoting or inhibiting rest and recreation. First, 18 items were drafted as candidate questions concerning attributes thought to pertain to rest and recreation. These included marital status, living environment, and number of days off per year. Next, 13 questions asking whether participants were practicing items from “the Rest and Recreation Guidelines for Health Promotion”¹² published by the Ministry of

Health, Labour and Welfare in 1994 were also added as candidate questions addressing participants’ current rest and recreation status. Fifteen questions pertaining to factors thought to inhibit rest and relaxation were drafted using the Delphi method within the group of researchers. These 46 items were used for the RRQ-Scale. A previous study¹⁰ surveyed 400 members of the general population living in the Tokyo area and conducted an exploratory analysis of the conceptual factor structure of the 46 items, which ultimately led to the selection of 16 questions as observed variables of the model.

This previous study¹⁰ reported satisfactory goodness of fit for the constructed model [goodness of fit index (GFI) = 0.943]. The 16 items selected in the previous study¹⁰ were (1) stressful experiences, (2) leisure and travel, (3) free time, (4) easy-going lifestyle, (5) tranquil environment, (6) easy-going eating habits, (7) tranquility in nature, (8) time to rest, (9) places to go on days off, (10) scrutiny from others, (11) friends and peers, (12) interest in rest, (13) experience and technique, (14) how to rest, (15) effects of rest, and (16) need for rest. In the present study, a team of experts reexamined the questions and construct model of the previous study.¹⁰ Reexamining these 16 items from the perspective of the factor structure demonstrated in the previous study,¹⁰ while considering construct validity and today’s social climate, resulted in the removal of four questions: “stressful experiences,” “leisure and travel,” “free time,” and “need for rest.” Consequently, 12 of the previous study’s questions were selected for use in this study.¹⁰ Table 1 shows the questionnaire used in the present study. The questionnaire is divided into two sections – Question 1 and Question 2 – consisting of four and eight questions, respectively. Questions 1.1–1.4 are answered using a 4-point scale (1 = Never, 2 = Rarely, 3 = Sometimes, 4 = Often). Questions 2.1–2.8 are answered using a different 4-point scale (1 = Strongly disagree, 2 = Disagree, 3 = Agree, 4 = Strongly agree). The total score (range 12–48 points) is calculated by adding all of the raw scores. The higher the score, the better the quality of rest/recreation.

Participants and survey method

In the present study, the RRQ-Scale was evaluated for floor and ceiling effects,¹² factorial validity,¹³ criterion-related validity,¹⁴ and reliability (Cronbach’s coefficient α ¹⁵) to consider whether it had sufficient practicality. We estimated the survey sample size needed to achieve this objective was about a minimum of 100 people, based on the recommendations of Terwee et al.¹³ With this estimate in mind, we asked 26 manufacturing-related

Table 1. The rest/recreation on day off quality scale for workers

*Question 1: For the following questions, please choose the one that is closest to your experience in everyday life or to your attitude toward life.

- 1 I use my free time to find meaning in life through hobbies and amusement.
- 2 I find ways to feel comfortable and at ease, even in familiar environments.
- 3 I find ways to take pleasure in food.
- 4 I try to obtain enjoyment and relaxation by being in nature, playing sports, or engaging in other leisure activities.

†Question 2: For the following questions on rest/recreation, choose the one that best describes your situation.

- 1 I do not have any places I can visit on my days off.
- 2 I do not have any acquaintances or friends to do things with on my days off.
- 3 Even when I take time off, I am not interested in doing anything.
- 4 Even if I wanted to do something on my days off, I have no experience or skills related to hobbies or other pursuits.
- 5 I do not really understand how to relax.
- 6 Taking time off is not very effective for reducing mental fatigue.
- 7 I'm too busy to take time off.
- 8 I do not feel like I can relax, because I worry about what other people will think.

*Questions 1.1–1.4 have 4 possible responses: Never, Rarely, Sometimes, Often. Responses are scored as "Never" = 1 point to "Often" = 4 points. †Questions 2.1–2.8 have 4 possible responses: Strongly disagree, Disagree, Agree, Strongly agree. Responses are scored as "Strongly disagree" = 1 point to "Strongly agree" = 4 points.

companies located in Oita Prefecture, Japan to participate in the study and surveyed all 796 employees of these firms. The participating firms' industries can be broken down as follows: transportation ($n = 2$), factory operations-related ($n = 3$), machine manufacturing-related ($n = 13$), power-related ($n = 3$), construction-related ($n = 4$), and administrative ($n = 1$). The minimum number of employees at one company was 3, and the maximum was 259. We decided these were suitable participants for this study because they were regular workers this scale could potentially be used on. For this study, self-administered questionnaires were individually distributed to subjects. With consideration to privacy, participants completed the questionnaires individually and returned them in unsigned envelopes. Concerning the method of questionnaire distribution, questionnaires for all study participants were first delivered to the general office managing the company as a whole. Subsequently, this office delivered the appropriate number of questionnaires to each participating firm, where they were individually distributed to employees. Regarding questionnaire collection, the envelopes containing the questionnaires submitted by individual employees were first collected at each firm and then delivered to the general office. When responses from all companies were present, they were delivered from the general office to the research institution. Self-administered questionnaires were distributed in February and March of 2017 and collected later.

Ethical considerations

The participants provided informed consent after receiving oral and written explanations of the purpose of the study, its data processing methods, and that they could refuse to participate or drop out without consequences. This study was conducted with the approval of the occupational health research ethics review committee of the Japan Organization of Occupational Health and Safety (notification number 18: March 25, 2016).

Data analysis

Calculation of descriptive statistics

First, the score distributions for each question were examined. After calculating the mean and SD of the scores for each question, we investigated whether the mean \pm SD were within the possible range of 1–4 points to check each question for floor and ceiling effects. This analysis was performed for three groups: all workers, only blue-collar workers, and only white-collar workers. Next, we calculated total scores for each sex to examine sex differences. This was done using the Mann-Whitney U test. Third, the total scores for different age classes were calculated to examine differences based on age. Three age classes were used: under 30 years old, 30–59 years old, and 60 years old and over. These were examined using the Kruskal-Wallis test. Further, multiple comparisons of score distribution were performed using the Steel-Dwass test for the total score of the RRQ-Scale for three groups: all workers, only blue-collar workers,

and only white-collar workers.

Examination of validity

Factorial validity; First, exploratory factor analysis¹⁴ was performed on the scores of all questions. Maximum-likelihood estimation was used as the method of factor extraction, and promax rotation was used as the rotation method. The number of factors was the number with an eigenvalue of 1 or more. Correlations between the extracted factors were also examined.

Next, the goodness of fit of the higher-order factor model was verified by confirmatory factor analysis.¹⁵ To ensure discrimination of the model, the pass value from each extracted factor to one question was constrained to 1. The pass values from error variances (e) were also constrained to 1. For fit indices, we evaluated the goodness of fit index (GFI), adjusted goodness of fit index (AGFI), comparative fit index (CFI), root mean square residual (RMR), and root mean square error of approximation (RMSEA).¹⁶ This analysis was performed for three groups: all workers, only blue-collar workers, and only white-collar workers.

Criterion-related validity (Concurrent validity); The criterion-related validity of the RRQ-Scale and SF-36v2®: MOS 36-Item Short-Form Health Survey version 2.0^{17–19} as health-related quality of life (HRQOL) were examined. SF-36v2® is the most commonly used scale for evaluating HRQOL in the world. Specifically, this study used the physical component summary (PCS), mental component summary (MCS), and role/social component summary (RCS), which are SF-36v2® summary scores. The factor coefficients used to calculate the summary scores were from the 2007 National Survey of Japan data.¹⁹ The standardized average of the national standard value is 50 points, and the standard deviation is 10 points, with higher scores being better HRQOL. Spearman's rank correlation coefficients were used to test how the questionnaire's total score and subscale scores correlated with PCS, MCS, and RCS. In addition, the same method was used to examine the criterion-related validity of the RRQ-Scale and the generalized self-efficacy scale. Self-efficacy is the perception of one's potential to successfully execute the behaviors required in a certain situation and is one of the core concepts of social learning theory.²⁰ Self-efficacy is divided into task-specific self-efficacy and generalized self-efficacy.^{21, 22} Task-specific self-efficacy is a form of self-efficacy that specifically impacts behavior regarding a task or situation, whereas generalized self-efficacy is independent of individual tasks or situations but impacts behavior in everyday situations in a more generalized and long-term manner.^{21, 22} The present

study used the Japanese version of the generalized self-efficacy scale,²³ which is a translation of the generalized self-efficacy scale developed by Sherer²² that has been tested for validity and reliability. This scale is composed of 23 questions and each question is assigned a score of 1 to 5 points, with total scores ranging from 23 to 115 points. Spearman's rank correlation coefficient was used to examine how the questionnaire's total score and subscale scores correlated with the generalized self-efficacy scale.

Examination of reliability

Cronbach's coefficient α was used to test the reliability of the questionnaire's total score and subscale scores. This analysis was performed for three groups: all workers, only blue-collar workers, and only white-collar workers. The statistical software IBM SPSS® Statistics version 22 (IBM, Armonk, NY) and IBM SPSS® Amos version 22 (IBM, Armonk, NY) were used for data analysis. The standard for statistical significance was set at $P < 0.05$. (two-sided test).

Setting cut-off points using methods based on characteristics of the distribution

The cut-off points for the questionnaire's total score were calculated using methods based on characteristics of the distribution.²⁴ First, the SE of the questionnaire's total score was calculated using the following formula: $SE = SD \times \sqrt{1 - \text{reliability}}$.²⁴ SD was calculated from sum of the questionnaire's total score, and reliability was calculated using Cronbach's coefficient α . In addition, based on the proposal of McHorney et al.,²⁵ cut-off points were set to 1.96 SE, which was equivalent to a statistically significant difference. Specifically, high quality was defined as mean total score +1.96 SE or higher, medium quality as mean -1.96 SE to mean +1.96 SE, and low quality as less than mean -1.96 SE.

RESULTS

Of the survey's 796 subjects, questionnaires were collected from 789 (response rate: 99.1%), of which 756 had valid responses to all 12 questions of the RRQ-Scale (effective response rate: 95.8%). Of the 764 subjects who had valid responses to all questions of SF-36v2® (effective response rate: 96.0%), 750 had valid responses to all questions on the Generalized Self-Efficacy Scale (effective response rate: 94.2%). Analyses for the present study used data from the 756 subjects who had valid responses to all questions of the RRQ-Scale. The data of these participants, which included 694 men (91.8%), 62 women (8.2%), 9 unknown (1.2%) were analyzed. Their average age was 44.7 ± 13.5 years within the

range of 18–81 years. Analysis of subject characteristics and score distributions for the SF-36v2[®] subscales (physical component score [PCS], mental component score [MCS], and role/social component score [RCS]) and the Generalized Self-Efficacy Scale are shown by worker type (blue-collar or white-collar) in Table 2. The subjects included 634 blue-collar workers and 78 white-collar workers. A total of 176 subjects were shift workers (23.3%). Scores (mean \pm SD) on the SF-36v2[®] (PCS, MCS, and RCS subscales) and self-efficacy scale for all workers were, respectively, 49.4 ± 10.3 , 47.6 ± 10.1 , 46.8 ± 12.0 , and 71.0 ± 11.0 .

Descriptive statistics of scale

Distribution of scores

Table 3 shows the questions (Q1.1–Q1.4 and Q2.1–Q2.8) and the distribution of total scores (all/blue-collar/white-collar workers). In all workers, the smallest mean \pm SD was Q1.4 at 2.53 ± 0.89 points and the largest was Q2.6 at 3.12 ± 0.85 points. All of the questions' mean \pm SD were within the maximum possible range (1–4 points, implying the absence of floor and ceiling effects). The mean \pm SD of the total score was 34.65 ± 6.26 points, and the distribution had kurtosis of -0.20 and skewness of -0.03 .

Gender or age differences in total scores

The mean \pm SD total score for men was 34.54 ± 6.23 points and 35.77 ± 6.57 points for women. The total scores for both sexes were not significantly different ($P = 0.238$). The mean \pm SD for subjects under 30-years-old was 35.53 ± 5.94 points, 30 to 59-years-old was 34.24 ± 6.32 points, and 60-years-old and older was 35.16 ± 6.12 points. Statistically significant differences in total score were not observed between age classes ($P = 0.114$).

Type of worker differences in total scores

In blue-collar workers, the mean \pm SD of the total score was 34.50 ± 6.27 points. In white-collar workers, the mean \pm SD of the total score was 35.73 ± 6.12 points. The results of multiple comparisons of the total score of the RRQ-Scale revealed no significant differences between any of the three groups (all and blue-collar, $P = 0.941$; all and white-collar workers, $P = 0.438$; blue-collar and white-collar workers, $P = 0.357$).

Validity of scale

Factorial validity

An exploratory factor analysis showed the changes in the eigenvalue of correlation matrix were 5.21, 1.68, 1.12, and 0.74, indicating that a 3-factor structure was

suitable. Table 4 shows the final factor patterns and factor correlations after promax rotation, assuming a 3-factor structure. The ratio explaining the total variance of the 12 items with the 3 factors before rotation was 55.94%. Factor I was composed of four items, with items on actively acquiring rest/recreation having a high load. These included “I use my free time to find meaning in life through hobbies and amusement” and “I find ways to take pleasure in food.” Therefore, Factor I was named “Rest and Recreation Acquisition.” Factor II was composed of six items, with items on the individual's ability to fully experience rest/recreation having a high load. These included “Even if I wanted to do something on my time off, I have no experience or skills related to hobbies or other pursuits” and “I don't really understand how to relax.” Therefore, Factor II was named “Rest and Recreation Ability.” Factor III was composed of two items, with items on the circumstances surrounding the individual having a high load. These included “I don't feel like I can relax because I worry about what other people will think.” Therefore, Factor III was named “Rest and Recreation Circumstances.”

Next, Fig. 1 shows the solutions (standardizing coefficients) of the confirmatory factor analysis (all/blue-collar/white collar). In all workers, the pass values from the observed variables (questions) were 0.71–0.82 to Factor I (first-order factor) 0.52–0.86 to Factor II, and 0.57–0.79 to Factor III, which are sufficiently large positive values. The pass values from the first-order factors to the higher-order factor were also sufficiently large positive values from 0.63 to 0.85. All parameters were statistically significant, except for those that were constrained ($P < 0.001$). The fit of the model was sufficiently high; GFI = 0.955, AGFI = 0.932, CFI = 0.962, RMR = 0.024, and RMSEA = 0.065.¹⁶ The GFI for the only blue-collar workers model was 0.953. The GFI for the only white-collar workers model was 0.810.

Criterion-related validity (Concurrent validity)

Table 5 shows the correlation coefficients between the RRQ-Scale and SF-36v2[®] or generalized self-efficacy scale. The correlations between the RRQ-Scale and PCS were $r_s = 0.193$ for total score ($P < 0.001$), 0.173 for factor I ($P < 0.001$), 0.151 for factor II ($P < 0.001$), and 0.098 for factor III ($P = 0.008$). Correlations with MCS were 0.369 for total score ($P < 0.001$), 0.288 for factor I ($P < 0.001$), 0.364 for factor II ($P < 0.001$), and 0.248 for factor III ($P < 0.001$). Correlations with RCS were 0.115 for total score ($P = 0.002$), 0.088 for factor I ($P = 0.017$), 0.029 for factor II ($P = 0.426$), and 0.211 for factor III ($P < 0.001$). Correlations between the RRQ-Scale and generalized self-efficacy scale were 0.397 for total score (P

Table 2. Characteristics of workers (Blue-collar/White-collar/All)

	Blue-collar (N = 634)	White-collar (N = 78)	Unknown (N = 44)	All workers (N = 756)
Sex				
Male	95.9%	52.6%	77.3%	90.3%
Female	3.3%	47.4%	6.8%	8.1%
Unknown	0.8%	0.0%	15.9%	1.6%
Age class				
< 30 yr	16.6%	10.3%	4.5%	15.2%
≥ 30 yr or < 60 yr	68.8%	71.8%	50.0%	68.0%
≥ 60 yr	14.4%	17.9%	25.0%	15.3%
Unknown	0.3%	0.0%	20.5%	1.5%
Job category				
Management	11.8%	38.5%	11.4%	14.6%
Non-management	77.8%	52.6%	11.4%	71.3%
Unknown	10.4%	9.0%	77.3%	14.2%
Shift work				
No	71.8%	93.6%	38.6%	72.1%
Yes	27.3%	1.3%	4.5%	23.3%
Unknown	0.9%	5.1%	56.8%	4.6%
Overtime per month				
< 45 hrs/month	83.1%	80.8%	63.6%	81.7%
≥ 45 hrs/month or < 60 hrs/month	6.2%	2.6%	9.1%	6.0%
≥ 60 hrs/month or < 80 hrs/month	6.0%	5.1%	15.9%	6.5%
≥ 80 hrs/month or < 100 hrs/month	0.6%	5.1%	0.0%	1.1%
≥ 100 hrs/month	1.3%	0.0%	4.5%	1.3%
Unknown	2.8%	6.4%	6.8%	3.4%
Number of days off per month				
< 4 days	10.3%	6.4%	11.4%	9.9%
4–7 days	46.5%	28.2%	45.5%	44.6%
8–11 days	38.8%	50.0%	18.2%	38.8%
≥ 12 days	3.0%	14.1%	18.2%	5.0%
Unknown	1.4%	1.3%	6.8%	1.7%
SF-36v2 score				
PCS				
< 25 points	2.2%	2.2%	0.0%	2.0%
≥ 25 points or < 50 points	43.4%	43.4%	50.0%	43.4%
≥ 50 points or < 75 points	52.5%	52.5%	47.7%	52.9%
≥ 75 points	0.2%	0.2%	0.0%	0.1%
Unknown	1.7%	1.7%	2.3%	1.6%
MCS				
< 25 points	2.4%	1.3%	2.3%	2.2%
≥ 25 points or < 50 points	55.8%	55.1%	50.0%	55.4%
≥ 50 points or < 75 points	40.1%	43.6%	43.2%	40.6%
≥ 75 points	0.0%	0.0%	2.3%	0.1%
Unknown	1.7%	0.0%	2.3%	1.6%
RCS				
< 25 points	4.6%	5.1%	6.8%	4.8%
≥ 25 points or < 50 points	46.5%	47.4%	52.3%	47.0%
≥ 50 points or < 75 points	47.2%	47.4%	38.6%	46.7%
≥ 75 points	0.0%	0.0%	0.0%	0.0%
Unknown	1.7%	0.0%	2.3%	1.6%
Self-efficacy scale score				
< 25 points	0.0%	0.0%	0.0%	0.0%
≥ 25 points or < 50 points	3.8%	5.1%	2.3%	3.8%
≥ 50 points or < 75 points	64.0%	48.7%	56.8%	62.0%
≥ 75 points	29.0%	43.6%	29.5%	30.6%
Unknown	3.2%	2.6%	11.4%	3.6%

hrs, hours; MCS, Mental Component Summary; PCS, Physical Component Summary; RCS, Role/Social Component Summary; SF-36v2, MOS 36-Item Short-Form Health Survey version 2.0; yr, year(s).

Table 3. Descriptive statistics for the responses to the twelve questions of the rest and recreation quality scale for workers (ceiling and floor effect, All/Blue-collar/White-collar)

All workers (<i>N</i> = 756)						
Question	Minimum score	Maximum score	Mean score	SD	Mean-SD	Mean+SD
Q1.1	1	4	2.76	0.79	1.97	3.55
Q1.2	1	4	2.67	0.74	1.93	3.41
Q1.3	1	4	2.73	0.75	1.98	3.48
Q1.4	1	4	2.53	0.89	1.64	3.42
Q2.1	1	4	2.88	0.81	2.07	3.69
Q2.2	1	4	2.87	0.88	1.99	3.75
Q2.3	1	4	2.92	0.83	2.09	3.75
Q2.4	1	4	3.01	0.82	2.19	3.83
Q2.5	1	4	3.21	0.74	2.47	3.95
Q2.6	1	4	3.12	0.85	2.27	3.97
Q2.7	1	4	2.80	0.72	2.08	3.52
Q2.8	1	4	3.13	0.77	2.37	3.90
Total scores of 12 questions	13	48	34.65	6.26	28.40	40.90
Only blue-collar workers (<i>N</i> = 634)						
Q1.1	1	4	2.75	0.79	1.96	3.54
Q1.2	1	4	2.67	0.75	1.93	3.42
Q1.3	1	4	2.72	0.76	1.96	3.48
Q1.4	1	4	2.52	0.87	1.65	3.39
Q2.1	1	4	2.82	0.70	2.12	3.52
Q2.2	1	4	2.87	0.81	2.06	3.68
Q2.3	1	4	3.12	0.77	2.35	3.89
Q2.4	1	4	2.87	0.89	1.98	3.76
Q2.5	1	4	2.90	0.84	2.06	3.74
Q2.6	1	4	2.99	0.83	2.16	3.82
Q2.7	1	4	3.19	0.75	2.44	3.94
Q2.8	1	4	3.08	0.87	2.21	3.95
Total scores of 12 questions	13	48	34.50	6.27	28.23	40.77
Only white-collar workers (<i>N</i> = 78)						
Q1.1	1	4	2.78	0.78	2.00	3.56
Q1.2	1	4	2.67	0.68	1.99	3.35
Q1.3	1	4	2.87	0.65	2.22	3.52
Q1.4	1	4	2.58	0.93	1.65	3.51
Q2.1	1	4	2.63	0.78	2.07	3.69
Q2.2	1	4	3.04	0.81	1.99	3.75
Q2.3	1	4	3.19	0.72	2.09	3.75
Q2.4	1	4	2.96	0.90	2.19	3.83
Q2.5	1	4	3.13	0.75	2.47	3.95
Q2.6	1	4	3.19	0.77	2.27	3.97
Q2.7	1	4	3.36	0.68	2.08	3.52
Q2.8	1	4	3.33	0.80	2.37	3.90
Total scores of 12 questions	22	48	35.73	6.12	29.61	41.85

N = 756. Extraction factors method: maximum likelihood method. Number of factors: number of correlation matrix eigenvalue is 1 or higher. Rotation method: promax rotation. The bold values indicate single-factor loading in the items is 0.350 or higher. Factor I, Rest and Recreation Acquisition; Factor II, Rest and Recreation Ability; Factor III, Rest and Recreation Circumstances.

Table 4. Factor analysis of rest/recreation quality scale for workers and its sub-scales (All workers)

Questions	Factor loading		
	Factor I	Factor II	Factor III
Q1.1	0.741	0.080	0.036
Q1.2	0.848	−0.033	−0.022
Q1.3	0.767	−0.064	0.058
Q1.4	0.702	0.046	−0.060
Q2.1	−0.016	0.668	0.085
Q2.2	−0.027	0.602	0.107
Q2.3	0.055	0.858	−0.041
Q2.4	0.025	0.945	−0.161
Q2.5	−0.064	0.812	0.044
Q2.6	0.103	0.365	0.176
Q2.7	0.038	−0.032	0.606
Q2.8	−0.038	0.098	0.705
Factor correlation			
Factor I	1.000	0.525	0.511
Factor II	0.525	1.000	0.397
Factor III	0.511	0.397	1.000

$N = 756$. Extraction factors method: maximum likelihood method. Number of factors: number of correlation matrix eigenvalue is 1 or higher. Rotation method: promax rotation. The bold values indicate single-factor loading in the items is 0.350 or higher. Factor I, Rest and Recreation Acquisition; Factor II, Rest and Recreation Ability; Factor III, Rest and Recreation Circumstances.

< 0.001), 0.363 for factor I ($P < 0.001$), 0.349 for factor II ($P < 0.001$), and 0.137 for factor III ($P < 0.001$).

Reliability of scale

In all workers, the reliability of all questions (Cronbach's coefficient α) was $\alpha = 0.877$. Factor I was $\alpha = 0.849$, Factor II was 0.870, and Factor III was 0.617. For blue-collar workers, Cronbach's α coefficient was 0.876. In white-collar workers, Cronbach's α coefficient was 0.883 (Table 6).

Cut off points of total score

The SE of the total score was 2.19. Therefore, the upper bound of the of cut-off points was 38.94 points ($34.65 + 1.96 \times 2.19$) and the lower bound was 30.36 points ($34.65 - 1.96 \times 2.19$). Considering these cut-off points, a total score of 12–30 points was defined as low quality, 31–38 points as medium quality, and 39–48 points as high quality.

DISCUSSION

The RRQ-Scale score distribution showed no floor or ceiling effects. Confirmatory factor analysis results verified the factorial validity. Consistent with the theoretical predictions, the total score of the RRQ-Scale

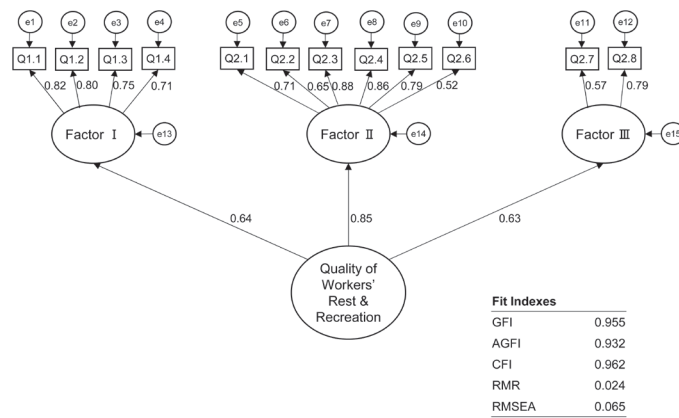
demonstrated statistically significant positive correlations with the SF-36v2[®] (PCS, MCS, and RCS) and the Generalized Self-Efficacy Scale, thereby confirming criterion-related validity. Furthermore, the overall reliability of the scale was high ($\alpha = 0.877$). The RRQ-Scale score distribution for workers demonstrated good validity and reliability, indicating high practicality. Moreover, there were no significant differences in analyses by sex, age class, or worker type (blue-collar or white-collar), suggesting that the RRQ-Scale is a universal scale.

Validity is a concept used to assess whether what a scale is attempting to measure is really being measured.²⁴ Validity is a multifaceted concept, but it can be broadly divided into three categories: content validity, construct validity, and criteria validity.²⁶

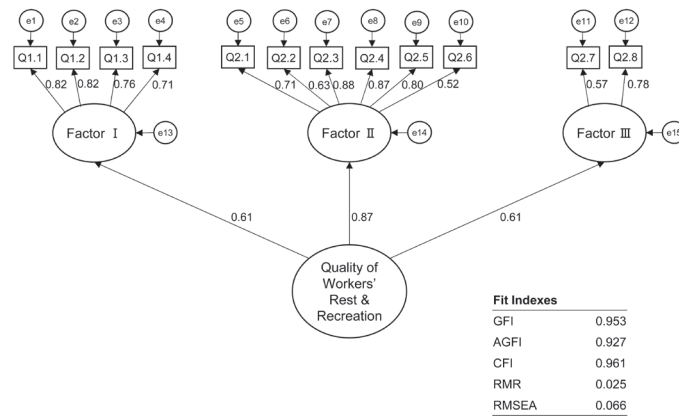
Content validity signals whether a scale's items reflect the content of the concept the scale is trying to measure without bias.²⁴ Usually, experts check the content of items created by referencing preliminary conceptual surveys and other results.²⁴ The questions of the RRQ-Scale were devised by a team of experts involved in creating the "Rest/Recreation Guideline for Health Promotion,"¹² a national guideline issued by the Japanese government. Consensus was reached through the Delphi method.²⁷ Of the 46 questions that were

Workers' rest and recreation quality scale on their days off

All workers (N = 736)



Only blue-collar workers (N = 634)



Only white-collar workers (N = 78)

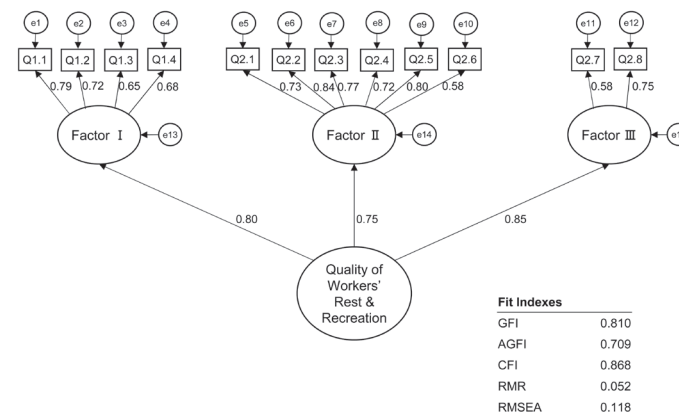


Fig. 1 . The result of confirmatory factor analysis (All/Blue-collar/White-collar). Factor I, Rest and Recreation Acquisition; Factor II, Rest and Recreation Ability; Factor III, Rest and Recreation Circumstances. GFI, Goodness of Fit Index; AGFI, Adjusted Goodness of Fit Index; CFI, Comparative Fit Index; RMR, Root Mean Square Residual; RMSEA, Root Mean Square Error of Approximation.

Table 5. Correlation coefficients between the rest and recreation quality scale and QOL scale (SF-36v2) and the generalized self-efficacy scale (All workers)

		SF-36v2®			Generalized self-efficacy scale
		PCS	MCS	RCS	
Sub-scales	Factor I	0.173**	0.288**	0.088*	0.363**
	Factor II	0.151**	0.364**	0.029	0.349**
	Factor III	0.098**	0.248**	0.211**	0.137**
Overall-scale	Scale total, 12 questions	0.193**	0.369**	0.115**	0.397**

N = 756. All scores were calculated by Spearman's rank correlation coefficient. Missing values were excluded. **P* < 0.05, ***P* < 0.01.

created, 16 were selected through factor analysis.¹⁰ At the outset of the present study, the expert team reexamined the questions from the previous study¹⁰ to select 12 final questions. Therefore, we believe the RRQ-Scale has sufficient content validity.

Factorial validity is a form of construct validity obtained when factor analysis confirms that a scale measuring a certain construct is composed of multiple sub-concepts.²⁸ In the present study, we first performed exploratory factor analysis¹⁴ to identify three factors, and then we examined their goodness of fit using confirmatory factor analysis.¹⁵ The overall factor structure of the RRQ-Scale exhibited sufficient goodness of fit; GFI = 0.955, indicating high factorial validity.¹⁶ The total scores were calculated by adding the scores of the sub-scales.²⁹

Criterion-related validity is examined regarding how a scale score relates to other similar scale scores.³⁰ A requirement of concurrent validity, which is a type of criterion-related validity, is that the scale used as the criterion has high reliability and validity, as well as measuring the same construct.³¹ In the present study, we first examined the concurrent validity of the RRQ-Scale and SF-36v2® (specifically, the PCS, MCS, and RCS sub-scales), which is a representative scale of HRQOL. Originally, we wanted to use another scale that measures worker rest/recreation as the external criterion of the RRQ-Scale, but no other scale currently satisfies the requirements. Therefore, as the main purpose of obtaining high-quality rest/recreation is to improve workers' mental and physical health and quality of life (QOL), we decided to use SF-36v2® as the external criterion, because this is an internationally standardized scale for measuring health and QOL.¹⁹ The correlations between RRQ-Scale total score and PCS/MCS/RCS were significant: *rs* = 0.193, 0.369, and 0.115, respectively. This demonstrates the relationships between HRQOL and the RRQ-Scale. Specifically, correlations with HRQOL subscales, PCS and RCS, were significant but extremely weak, while the correlation with the MCS was

somewhat higher compared to the other two subscales. This suggests that the constructs of the RRQ-Scale are independent of the physical and role/social aspects of HRQOL, overlapping only slightly with them, although there is greater overlap with the mental aspect. Second, we examined the concurrent validity of the RRQ-Scale and the generalized self-efficacy scale.³² Self-efficacy is an important psychological factor involved in selecting health behaviors that improve health. Acquiring good rest/recreation is considered a health behavior; thus, we hypothesized that high self-efficacy would lead to more implementation of health behaviors, such as acquiring good rest/recreation. Based on this hypothesis, we adopted the self-efficacy scale as a criterion. Consistent with prior predictions, our results showed that the correlation between the RRQ-Scale total score and the generalized self-efficacy scale was significant, with *rs* = 0.397 (*P* < 0.001). Considering these results on content validity, construct validity, and criterion validity, we believe the RRQ-Scale has sufficiently high validity.

Reliability of scale is a concept that indicates the degree of agreement between measurements of the same subject performed under different conditions.²⁴ Cronbach's α is the lower-limit value of reliability coefficient estimates and is considered desirable in psychological scales α = 0.7 or higher.^{31, 33} For the RRQ-Scale total score, the alpha was 0.877, satisfying the standards estimated for such scales. Factors I and II had good reliabilities; their alpha coefficients were α = 0.849 and 0.870, respectively. However, Factor III had a coefficient lower than 0.7 at α = 0.617. Cronbach's α tends to increase when there are many questions.²⁴ Factor III has only two questions, which may explain its lower α coefficient. Further, as Factor III's α coefficient was not much lower than 0.7, we believe it is within the acceptable range. For the present scale, the corrected item-total correlations between all of the questions and sub-scales (Factors I to III) were 0.356 to 0.801. Based on the above, we believe the RRQ-Scale has adequate breadth of content, internal consistency, and sufficient

Table 6. Reliability analysis of rest and recreation quality scale for worker (overall and sub-scales, All/Blue-collar/White-collar)

All workers (<i>N</i> = 756)							
Overall-scale	Questions	Corrected item-total correlation	Cronbach's coefficient α	Sub-scales	Questions	Corrected item-total correlation	Cronbach's coefficient α
Quality of rest/recreation for worker	Q1.1	0.637	0.877	Factor I	Q1.1	0.724	0.849
	Q1.2	0.567			Q1.2	0.721	
	Q1.3	0.538			Q1.3	0.673	
	Q1.4	0.509			Q1.4	0.648	
	Q2.1	0.609		Factor II	Q2.1	0.648	0.870
	Q2.2	0.558			Q2.2	0.606	
	Q2.3	0.743			Q2.3	0.801	
	Q2.4	0.698			Q2.4	0.773	
	Q2.5	0.666		Factor III	Q2.5	0.745	0.617
	Q2.6	0.500			Q2.6	0.475	
	Q2.7	0.356			Q2.7	0.447	
	Q2.8	0.464			Q2.8	0.447	
Only blue-collar workers (<i>N</i> = 634)							
Quality of rest/recreation for worker	Q1.1	0.631	0.876	Factor I	Q1.1	0.731	0.858
	Q1.2	0.567			Q1.2	0.734	
	Q1.3	0.542			Q1.3	0.688	
	Q1.4	0.508			Q1.4	0.654	
	Q2.1	0.346		Factor II	Q2.1	0.648	0.872
	Q2.2	0.610			Q2.2	0.606	
	Q2.3	0.445			Q2.3	0.801	
	Q2.4	0.542			Q2.4	0.773	
	Q2.5	0.744		Factor III	Q2.5	0.745	0.616
	Q2.6	0.710			Q2.6	0.475	
	Q2.7	0.669			Q2.7	0.445	
	Q2.8	0.506			Q2.8	0.445	
Only white-collar workers (<i>N</i> = 78)							
Quality of rest/recreation for worker	Q1.1	0.644	0.883	Factor I	Q1.1	0.658	0.804
	Q1.2	0.591			Q1.2	0.620	
	Q1.3	0.555			Q1.3	0.572	
	Q1.4	0.465			Q1.4	0.623	
	Q2.1	0.389		Factor II	Q2.1	0.651	0.875
	Q2.2	0.627			Q2.2	0.779	
	Q2.3	0.577			Q2.3	0.713	
	Q2.4	0.670			Q2.4	0.660	
	Q2.5	0.697		Factor III	Q2.5	0.771	0.603
	Q2.6	0.603			Q2.6	0.503	
	Q2.7	0.647			Q2.7	0.431	
	Q2.8	0.511			Q2.8	0.431	

reliability.

This study also verified whether the RRQ-Scale had the versatility to be used with different types of workers. Specifically, subjects were divided into blue-collar and white-collar workers to perform total score distribution analysis, reliability testing, and confirmatory factor analysis by worker type. The results found no significant differences due to worker type, suggesting that the RRQ-Scale is a versatile scale that can be used for any type of worker.

To the best of our knowledge, the RRQ-Scale is the first scale in the world developed to evaluate the quality of worker rest/recreation, an important element of non-working hours. Having only 12 items, a practical feature of this scale is its simplicity and length. When questionnaires are used in epidemiological surveys of regular workers, having a small number of questions is important to reduce the burden on and costs to the participants. A small number of questions also means it is easy for subjects to respond to other questions to evaluate items other than the quality of rest/recreation during the same survey. For example, a worker's entire life could effectively be evaluated on the limited space of a questionnaire by combining this scale with others that evaluate work or sleep.

The results of this study were obtained from workers in a specific region of Japan and show the distribution, validity, and reliability of scores in this range. Specifically, this research is limited in that subjects included only workers at manufacturing-related firms and thus did not cover a wide range of occupations or worker types. Other populations may exhibit differing results. Going forward, studies should be conducted on populations of various backgrounds, and those randomly selected to be more representative. A novel scale must have its various psychometric properties examined. The present study is a first step in this process. Repeated epidemiological studies using this scale could analyze the effects of rest/recreation quality on the mental and physical health of workers and the relationship between worker characteristics and quality of rest/recreation. This could also help create recommendations on how workers should spend their rest/recreation, based on their background. Thus far, days off have been managed mainly in terms of quantity, that is time off according to the number of days. However, research using this scale could generate findings on how to manage worker days off based on the standpoint of rest/recreation quality. In 2019, Japan launched a national policy – “Work System Reform¹¹⁾” aimed at making major changes for all workers. In these circumstances, a rest/recreation scale could be an extremely important tool. In the future, more

empirical research using this scale must be accumulated. In sum, we developed a psychometrically sound scale to assess the quality of rest and recreation among the working population.

Acknowledgments: We thank Shinji Akamine (Japan Oita Labor Bureau) for his cooperation with this research, and Ken'ichi Narita (Kwansei Gakuin University) for the license of the Japanese version of the generalized self-efficacy scale.

This research was supported by Industrial Health Research Grants from Japan Organization of Occupational Health and Safety and JSPS KAKENHI Grant Number 18K10067.

The authors declare no conflict of interest.

REFERENCES

- 1 van de Vijfeijke H, Leijten FRM, Ybema JF, van den Heuvel SG, Robroek SJW, van der Beek AJ, et al. Differential effects of mental and physical health and coping style on work ability: a 1-year follow-up study among aging workers. *J Occup Environ Med.* 2013;55:1238-43. DOI: [10.1097/JOM.0b013e3182a2a5e1](https://doi.org/10.1097/JOM.0b013e3182a2a5e1), PMID: [24064781](https://pubmed.ncbi.nlm.nih.gov/24064781/)
- 2 Schultz AB, Edington DW. Employee health and presenteeism: a systematic review. *J Occup Rehabil.* 2007;17:547-79. DOI: [10.1007/s10926-007-9096-x](https://doi.org/10.1007/s10926-007-9096-x), PMID: [17653835](https://pubmed.ncbi.nlm.nih.gov/17653835/)
- 3 Leijten FRM, van den Heuvel SG, Ybema JF, Robroek SJW, Burdorf A. Do work factors modify the association between chronic health problems and sickness absence among older employees? *Scand J Work Environ Health.* 2013;39:477-85. DOI: [10.5271/sjweh.3353](https://doi.org/10.5271/sjweh.3353), PMID: [23440271](https://pubmed.ncbi.nlm.nih.gov/23440271/)
- 4 Schuring M, Robroek SJW, Otten FWJ, Arts CH, Burdorf A. The effect of ill health and socioeconomic status on labor force exit and re-employment: a prospective study with ten years follow-up in the Netherlands. *Scand J Work Environ Health.* 2013;39:134-43. DOI: [10.5271/sjweh.3321](https://doi.org/10.5271/sjweh.3321), PMID: [22961587](https://pubmed.ncbi.nlm.nih.gov/22961587/)
- 5 Hummer J, Sherman B, Quinn N. Present and unaccounted for. *Occup Health Saf.* 2002;71:40-2, 4, 100, PMID: [11977562](https://pubmed.ncbi.nlm.nih.gov/11977562/)
- 6 Hansen CD, Andersen JH. Going ill to work – What personal circumstances, attitudes and work-related factors are associated with sickness presenteeism? *Soc Sci Med.* 2008;67:956-64. DOI: [10.1016/j.socscimed.2008.05.022](https://doi.org/10.1016/j.socscimed.2008.05.022), PMID: [18571821](https://pubmed.ncbi.nlm.nih.gov/18571821/)
- 7 Miraglia M, Johns G. Going to work ill: A meta-analysis of the correlates of presenteeism and a dual-path model. *J Occup Health Psychol.* 2016;21:261-83. DOI: [10.1037/ocp0000015](https://doi.org/10.1037/ocp0000015), PMID: [26550958](https://pubmed.ncbi.nlm.nih.gov/26550958/)
- 8 Kinman G, Wray S. Presenteeism in academic employees—occupational and individual factors. *Occup Med (Chic Ill).* 2018;68:46-50. DOI: [10.1093/occmed/kqx191](https://doi.org/10.1093/occmed/kqx191), PMID: [29351632](https://pubmed.ncbi.nlm.nih.gov/29351632/)
- 9 Guertler D, Vandelanotte C, Short C, Alley S, Schoeppe S, Duncan MJ. The association between physical activity, sitting time, sleep duration, and sleep quality as correlates of presenteeism. *J Occup Environ Med.* 2015;57:321-8. DOI: [10.1097/JOM.0000000000000355](https://doi.org/10.1097/JOM.0000000000000355), PMID: [25742538](https://pubmed.ncbi.nlm.nih.gov/25742538/)
- 10 Harano S, Miyake T, Yokoyama E, Suzuki K, Ohida T, Nozaki S. Development of the resting promotion-inhibition model for health promotion. *Health Sciences.* 2004;20:255-64. Japanese.

- 11 Ministry of Health Labour and Welfare. The 13th Occupational Safety & Health Program [Internet]. Japan; 2018. [cited 2020 Dec 18]. Available from: <https://www.mhlw.go.jp/content/11200000/000341159.pdf>
- 12 Itani T, Kaibara J, Ohkuma Y, Ohta T, Tsutui S, Nozaki M, et al. Rest/Recreation Guideline for Health Promotion. *New Diet Therapy*. 1994;10:78-80. Japanese.
- 13 Terwee CB, Bot SDM, de Boer MR, van der Windt DAWM, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol*. 2007;60:34-42. DOI: [10.1016/j.jclinepi.2006.03.012](https://doi.org/10.1016/j.jclinepi.2006.03.012), PMID: [17161752](https://pubmed.ncbi.nlm.nih.gov/17161752/)
- 14 Fabrigar LR, Wegener DT, MacCallum RC, Strahan EJ. Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*. 1999;4:272-99. DOI: [10.1037/1082-989x.4.3.272](https://doi.org/10.1037/1082-989x.4.3.272)
- 15 Thompson B. Exploratory and confirmatory factor analysis: Understanding concepts and applications. Washington, DC: American Psychological Association; 2004.
- 16 Hu LT, Bentler PM. Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria Versus New Alternatives. *Structural Equation Modeling-a Multidisciplinary Journal*. 1999;6:1-55. DOI: [10.1080/10705519909540118](https://doi.org/10.1080/10705519909540118)
- 17 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;51:1037-44. DOI: [10.1016/S0895-4356\(98\)00095-X](https://doi.org/10.1016/S0895-4356(98)00095-X), PMID: [9817121](https://pubmed.ncbi.nlm.nih.gov/9817121/)
- 18 Fukuhara S, Ware JE Jr, Kosinski M, Wada S, Gandek B. Psychometric and clinical tests of validity of the Japanese SF-36 Health Survey. *J Clin Epidemiol*. 1998;51:1045-53. DOI: [10.1016/S0895-4356\(98\)00096-1](https://doi.org/10.1016/S0895-4356(98)00096-1), PMID: [9817122](https://pubmed.ncbi.nlm.nih.gov/9817122/)
- 19 Fukuhara S, Suzukamo Y. Manual of SF-36v2 Japanese version. Kyoto, Japan: iHope International Inc.; 2015. Japanese.
- 20 Bandura A. Self-Efficacy Mechanism in Human Agency. *American Psychologist*. 1982;37:122-47. DOI: [10.1037/0003-066x.37.2.122](https://doi.org/10.1037/0003-066x.37.2.122)
- 21 Bandura A. Self-efficacy: toward a unifying theory of behavioral change. *Psychol Rev*. 1977;84:191-215. DOI: [10.1037/0033-295X.84.2.191](https://doi.org/10.1037/0033-295X.84.2.191), PMID: [847061](https://pubmed.ncbi.nlm.nih.gov/847061/)
- 22 Sherer M, Maddux JE, Mercandante B, Prentice-Dunn S, Jacobs B, Rogers RW. The Self-Efficacy Scale: Construction and Validation. *Psychological Reports*. 2016;51:663-71. DOI: [10.2466/pr0.1982.51.2.663](https://doi.org/10.2466/pr0.1982.51.2.663)
- 23 Narita K, Shimonaka Y, Nakazato K, Kawaai C, Sato S, Osada Y. A Japanese Version of the Generalized Self-Efficacy Scale - Scale Utility from the Life-Span Perspective. *Japanese Journal of Educational Psychology*. 1995;43:306-14. DOI: [10.5926/jjep1953.43.3_306](https://doi.org/10.5926/jjep1953.43.3_306) Japanese.
- 24 Streiner DL, Norman GR, Cairney J. Health measurement scales: A practical guide to their development and use, 5th ed. NY, US: Oxford University Press; 2015.
- 25 McHorney CA, Tarlov AR. Individual-patient monitoring in clinical practice: are available health status surveys adequate? *Qual Life Res*. 1995;4:293-307. DOI: [10.1007/BF01593882](https://doi.org/10.1007/BF01593882), PMID: [7550178](https://pubmed.ncbi.nlm.nih.gov/7550178/)
- 26 Landy FJ. Stamp Collecting Versus Science - Validation as Hypothesis-Testing. *American Psychologist*. 1986;41:1183-92. DOI: [10.1037/0003-066x.41.11.1183](https://doi.org/10.1037/0003-066x.41.11.1183)
- 27 Skulmoski GJ, Hartman FT, Krahn J. The Delphi Method for Graduate Research. *J Inf Technol Educ*. 2007;6:1-21. DOI: [10.28945/199](https://doi.org/10.28945/199)
- 28 Cohen RJ, Swerdlik ME, Phillips SM. Psychological testing and assessment: An introduction to tests and measurement: Mayfield Publishing Co; 1996.
- 29 Liang J, Bollen KA. The structure of the Philadelphia Geriatric Center Morale scale: a reinterpretation. *J Gerontol*. 1983;38:181-9. DOI: [10.1093/geronj/38.2.181](https://doi.org/10.1093/geronj/38.2.181), PMID: [6827034](https://pubmed.ncbi.nlm.nih.gov/6827034/)
- 30 Kane MT. Current concerns in validity theory. *Journal of Educational Measurement*. 2001;38:319-42. DOI: [10.1111/j.1745-3984.2001.tb01130.x](https://doi.org/10.1111/j.1745-3984.2001.tb01130.x)
- 31 Salkind N. Encyclopedia of Research Design. Thousand Oaks, CA: SAGE Publications, Inc.; 2010.
- 32 Conner M, Norman P. Predicting health behaviour. 2nd ed. Buckingham, England: Open University Press; 2005.
- 33 Kaplan RM, Saccuzzo DP. Psychological testing: Principles, applications, and issues. 4th ed. Belmont, CA, US: Thomson Brooks/Cole Publishing Co.; 1997.