Brazilian Journal of Physical Therapy 26 (2022) 100465

Brazilian Journal of Physical Therapy

https://www.journals.elsevier.com/brazilian-journal-of-physical-therapy

ORIGINAL RESEARCH

Brazilian version of need for recovery scale: Assessment of structural validity, criterion validity, and internal consistency

Luiz Augusto Brusaca[a], Cristiane Shinohara Moriguchi[b], Dechristian França Barbieri[a,c], Matthew Leigh Stevens[d], Ana Beatriz Oliveira[a,*]

[a] *Laboratory of Clinical and Occupational Kinesiology (LACO), Department of Physical Therapy, Universidade Federal de São Carlos, São Carlos, SP, Brazil*
[b] *Laboratory of Preventive Physical Therapy and Ergonomics (LAFIPE), Department of Physical Therapy, Universidade Federal de São Carlos, São Carlos, SP, Brazil*
[c] *Department of Industrial Engineering, Clemson University, Clemson, SC, United State*
[d] *National Research Centre for the Working Environment, Copenhagen, Denmark*

Received 16 November 2021; received in revised form 19 October 2022; accepted 21 November 2022
Available online 29 November 2022

KEYWORDS
COSMIN;
Measurement properties;
Need for recovery;
Occupational health;
Sustainable employment;
Workers

Abstract
Background: The Need For Recovery scale (NFR) is a tool that allows early identification of work-related health risks. However, the structure of the Brazilian version of NFR scale (Br-NFR) which contains 11 items has not been evaluated.
Objectives: To evaluate the structural validity, criterion validity, and internal consistency of the Br-NFR scale in workers.
Methods: 672 workers were included in this study. A split-half validation method was applied to the sample to create a development and validation sample. The structure of the Br-NFR was examined through Exploratory Factor Analysis (EFA) using the development sample. The validation sample was used to evaluate the structure with Confirmatory Factor Analysis (CFA). For the latter, several goodness-of-fit indices were considered to evaluate the model fit of the structures tested in this study. Criterion validity was assessed between the Brazilian structure and structures found in the literature compared with the original scale through intraclass correlation coefficient ($ICC_{2,1}$). The internal consistency of the Br-NFR was assessed using Cronbach's alpha. Both analyses used the validation sample.
Results: The EFA showed that the scale has a one-factor structure and the CFA demonstrated that the Br-NFR structure with 7 items presented excellent to acceptable goodness-of-fit indices. Excellent values of ICC were found between the structures tested in the study and the original 11-item structure of the NFR. The Br-NFR scale presented good internal consistency.

* Corresponding author at: Laboratory of Clinical and Occupational Kinesiology (LACO), Department of Physical Therapy, Universidade Federal de São Carlos (UFSCar), Washington Luiz Road, km 235, SP310, 13565-905, São Carlos, SP, Brazil.
E-mail: biaoliveira@ufscar.br (A.B. Oliveira).

https://doi.org/10.1016/j.bjpt.2022.100465

Conclusion: The Br-NFR is unidimensional. The final 7-item version presented to be equivalent to the original 11-item scale and also has good internal consistency.

Introduction

The maintenance of workers' health is a public health issue and a principle of sustainable development. The maintenance of a productive workforce is vital for a functional society. Therefore, the identification of work-related health risks is a primary purpose of occupational health services.[1,2] The Need For Recovery scale (NFR) is a tool that allows early identification of work-related health risks based on feelings of overload, irritability, lack of energy for new efforts, reduced performance, and social withdrawal.[3] The concept of NFR is based on a daily process of being able to recover from short-term fatigue symptoms, considering recovery opportunities during and after work. If recovery is unable to reverse fatigue symptoms, residual symptoms will remain the next day and an accumulation process can take place.[4] As such, the lack of recovery after a workday has been suggested to be an early indicator of sleep problems, depression, sickness absence, and psychosomatic complaints.[4-6] Thus, NFR is a valuable source of information for occupational health practitioners such as physicians, nurses, physical therapists, and researchers.

The NFR scale is an 11-item self-report instrument with valid measurement properties with regard to internal consistency (Cronbach's alpha = 0.81 – 0.92), test-retest reliability (intraclass correlation coefficient [ICC] = 0.68 – 0.80), unidimensionality of the construct, and sensitivity to detecting change.[3,7-10] The scale was originally developed in the Netherlands as part of the Dutch Questionnaire on the Experience and Evaluation of Work[8] and has been translated and cross-culturally adapted in different countries, such as Brazil,[10] France,[11] Italy,[12] China,[13] Iran,[14] Denmark,[15] and Sweden.[16]

In Brazil, the NFR scale has been used by physical therapists and researchers in the field of ergonomics to assess need for recovery for several jobs (e.g., office workers, nurses, and blue-collar workers) .[17-19] Worldwide, beyond its use to assess work conditions in large cohort studies,[3,9] occupational health services are also using the NFR to assess workplace interventions, such as: worksite aerobic exercises intervention[20] and participatory physical and psychosocial workplace intervention.[21]

The use of health measurement instruments for both clinical and research applications requires the assessment of their psychometric qualities for specific criteria.[22] Among these criteria, structural validity is of primary importance because it indicates whether the instrument adequately reflects the dimensionality of the construct to be measured.[23,24] Therefore, the structural validity defines the configuration of items to represent the intended construct and prevent the burden of unnecessary questions.[22-24]

The COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN group)[23,25] recommend that structural validity should be assessed by Confirmatory Factor Analysis (CFA), which is a statistical method of the Classic Test Theory (CTT) or by Item Response Theory. Moreover, COSMIN recommends applying an Exploratory Factor Analysis (EFA) while an instrument is being developed; in translation and cross-cultural adaptation studies; or when there are no previous hypothesis of the number of dimensions of an instrument.[23] Thus, EFA and CFA are used to reduce the number of items on a questionnaire, because items that have no contribution or have an unclear contribution to the factor structure can be deleted based on pre-defined criteria.[23,26,27] Furthermore, regardless of the measurement theory used, it is important to maintain precision and validity of the instrument to continue measuring the entire continuum (i.e., latent trait) of the health outcome without any loss.[23]

Regarding the NFR, most studies[11-13,15,16] only examined the factor structure (i.e., unidimensionality) of the scale through EFA, and solely the Chinese study[13,28] confirmed the structure of the scale using CFA. Even though there is a previous hypothesis of unidimensionality of the NFR scale,[3,9,13,28] a recent study conducted by Stevens et al.[15] (who developed and validated a short-form version of the Danish NFR) showed, through the content (face) validity assessment, that the items of the scale suggested a two-factor structure, i.e., one factor of 'recovery of mental resources' and another of 'recovery of physical resources'. However, after applying the EFA, it was possible to observe that the scale is unidimensional, which is in line with previous literature.[3,9,13,28]

The structural validity of the Brazilian version of NFR (Br-NFR) was not evaluated during the cross-cultural adaptation of the scale to Brazilian Portuguese.[10] Therefore, the main aim of the present study was to evaluate, through CTT, the structural validity, criterion validity, and internal consistency of the Br-NFR scale. Our hypothesis is that the Br-NFR will present an unidimensional structure consistent with the original version[3,9] and also recent studies.[13,15,28] We also hypothesize that when comparing the structures of other language versions with the structure from this study, ours will show acceptable results, with excellent goodness-of-fit indices (i.e., values $\langle$ or $\rangle$ than the cut-off points), excellent criterion validity (i.e., ICC $\geq$0.95), and acceptable internal consistency (i.e., Cronbach's alpha $>$0.70).

Methods

Design

This study comprised a study of measurement properties using cross-sectional data. The measurement theories and analysis used in this study follow the guidelines outlined by Costello and Osborne,[27] Prinsen et al.[25] (COSMIN group), and Worthington and Whittaker.[26]

Participants

The sample consists of 672 workers from three previous studies[10,29] – 192 blue-collar workers from the manufacturing industry (n = 141 females and n = 51 males, mean ± standard deviation (SD) age 34.5 ± 8.3 years);[10] 370 call center operators from a private electrical company (n = 168 females and n = 202 males, mean age 26.8 ± 7.7 years);[29] and 110 office workers from a public university in Brazil (n = 63 females and n = 47 males, mean age 33.8 ± 9.5 years; unpublished data). The studies were conducted in accordance with the Helsinki declaration, and all participants provided their written informed consent prior to entering in each study. The Human Ethics Committee of the Universidade Federal de São Carlos approved the studies by the registration processes #0054.0.135.000.07, #1080.0.000.135–10, and #0068.0.135.000–10.

Outcomes and data collection

Collected outcome measures used in this study include the Br-NFR scale[10] and basic demographic information. The Br-NFR scale is an 11-item Likert scale with four response categories: "Never" = 0; "Sometimes" = 1; "Often" = 2; and "Always" = 3. The individual sum of these scores is converted to an index from 1 to 100, where 100 indicates the maximum requirement for recovery. All items included in the Br-NFR scale and their English version are provided in Table 1.

Data analysis

For the analyses we used a split-half validation method in which the participants of these three groups were randomly divided into two groups – a development sample and a validation sample – each with 336 participants.[23,26] The random division was performed using the 'sample' function of the software R v4.1.1,[30] which allocates each worker to one of the two groups without using any external variables. Thus, avoiding introducing any bias in sample allocation. We then used the development sample to explore the factor structure using EFA. We used the validation sample to check the structural validity, criterion validity, and internal consistency of the structure identified by the EFA using CFA, ICC, and Cronbach's alpha.

Descriptive statistics

Descriptive data summaries are presented by means of absolute number [percentage] or mean ± standard deviation. In addition, the response distributions and the correlation matrix are shown for each item of the scale for the complete sample as descriptive data, as well as the missing responses for each sample.

Exploratory factor analysis

The EFA was used to examine the factor structure of the Br-NFR with the development sample. Suitability (factorability)

Table 1 Items of the "need for recovery scale" in Brazilian Portuguese and English.

Items	Brazilian version	English version
1	Eu acho difícil relaxar no fim de um dia de trabalho.	I find it difficult to relax at the end of a working day.
2	Ao fim do dia de trabalho eu me sinto realmente acabado (a).	By the end of the working day, I feel really worn out.
3	Por causa do meu trabalho, ao fim do dia eu me sinto muito cansado (a).	Because of my job, at the end of the working day I feel rather exhausted.
4	À noite, após um dia de trabalho, eu me sinto bem-disposto (a).	After the evening meal, I generally feel in good shape.
5	Eu precisode mais de um dia de folga do trabalho para começar a me sentir relaxado (a).	In general, I only start to feel relaxed on the second non-working day.
6	Eu acho difícil prestar atenção ou me concentrar durante meu tempo livre depois de um dia de trabalho.	I find it difficult to concentrate in my free time after work.
7	Eu acho difícil me interessar por outras pessoas assim que eu chego do trabalho.	I cannot really show any interest in other people when I have just come home myself.
8	Eu precisode mais de uma hora para me sentir completamente descansado (a) depois de um dia de trabalho.	Generally, I need more than an hour before I feel completely recuperated after work.
9	Quando eu chego em casa após o trabalho eu precisoser deixado em paz por um tempo.	When I get home from work, I need to be left in peace for a while.
10	Depois de um dia de trabalho eu me sinto tão cansado (a) que não consigo fazer outras atividades.	Often, after a day's work I feel so tired that I cannot get involved in other activities.
11	Na última parte do meu dia de trabalho, o cansaço me impede de fazer meu trabalho tão bem quanto eu normalmente faria se não estivesse cansado (a).	A feeling of tiredness prevents me from doing my work as well as I normally would during the last part of the working day.
	Categorias de Resposta	Response Categories
	1. Nunca	1. Never
	2. Algumas vezes	2. Sometimes
	3. Frequentemente	3. Often
	4. Sempre	4. Always

Note: For more information on the Brazilian and English version, see Moriguchi et al.[10] and van Veldhoven and Broersen.[3]

of the data set for EFA was evaluated based on Bartlett's test of sphericity – used to assess whether the correlation between items was adequate based on a criterion of $p<0.05$, and the Kaiser-Meyer-Olkin (KMO) statistic – used to measure sample adequacy based on a criterion of ≥ 0.80.[26]

Through EFA, the factor structure was explored with the implementation of a polychoric correlation matrix with the Minimum Rank Factor Analysis (MRFA) extraction method.[31] If a structure with more than one factor was found, the solutions for the factors would be examined using the direct Oblimin (oblique) rotation to ensure independence of the items, otherwise, the rotation would be set to none. We identified the number of factors to be retained through parallel analysis with random permutation of the observed data.[26,31]

To the best of our knowledge, there is no consensus on the criteria for item deletion or retention. Therefore, for a one-factor structure, items with a factor loading <0.32 and communality <0.40 were considered candidates for deletion.[26,27] On the other hand, for a structure with two- or more factors, items with factor loading <0.32; communalities <0.40; cross-loadings <0.15 in difference from the highest factor loading among an item; factor loading >0.32 in two or more factors were considered candidates for deletion.[26,27] It is noteworthy that care has been taken in applying all of these criteria, because EFA is an interactive, multi-step process. Moreover, the highest factor load was used to decide whether an item loaded on a specific factor.

Confirmatory factor analysis

The structural validity of the factor structure extracted during EFA was evaluated using CFA on the validation sample. The CFA was performed with the implementation of a polychoric matrix and the Robust Diagonally Weighted Least Squares (RDWLS) extraction method.[32] Considering that the Br-NFR is a Likert-type scale with four response categories, the use of RDWLS is more suitable for ordinal data than other extraction methods (e.g., maximum likelihood).[32] The model fit of the factor structure was evaluated using several indices: Pearson's chi-square (χ^2) test; chi-square divided by the degree of freedom (χ^2/df ratio; <3.00 is adequate, <5.00 is acceptable); Comparative Fit Index (CFI) and Tucker-Lewis Index (TLI; >0.90 is acceptable, >0.95 is excellent); the Root Mean Square Error of Approximation (RMSEA) with the 90% confidence interval (RMSEA; <0.08 is acceptable, <0.06 is excellent); and standardized root mean square residual (SRMR; <0.08 is acceptable).[26,33] Additionally, Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) were considered to compare the structures tested in this study (see the description below) and the lowest value was considered the most appropriate.[26,33] Moreover, factor loads ≥ 0.40 were considered adequate for the domain.

The structure extracted by EFA (i.e., Brazilian structure) was compared with five structures found in the literature: Dutch,[3,9] Swedish,[16] Iranian,[14] and Chinese[13] (11-item [Dutch being the original structure]); French[11] (10-item, item 10 was excluded); Italian[12] (10-item, item 4 was excluded); and Danish[15] (9-item, because items 2 and 3 were merged on the original scale, we compared two structures of the Danish scale with our own – one without items 2 and 11 [Danish 1] and one without items 3 and 11 [Danish 2]). All these structures contain the same concept with a one-factor structure that considers all items belonging to a domain of need for recovery. Moreover, the evaluation of the structures involved checking the quality of the goodness-of-fit indices. In addition, the standardized residual covariance matrix and the modification indices of the Brazilian structure were verified, then theoretical decisions were made to item deletion or retention, to identify which items would be more likely to represent the Br-NFR concepts.[26,33]

Criterion validity and internal consistency

The total scores of the previously specified structures (i.e., Brazilian structure validated through CFA, French, Italian, Danish 1 and 2) were compared with the total score of the original scale of 11-item using the $ICC_{2,1}$ (two-way random, single measurement, absolute agreement) to assess criterion validity.[23] For the purposes of validation, it was decided *a priori* that the criterion validity between the scales should be excellent (i.e., ICC ≥ 0.95). This comparison was performed to assess whether the structures are an adequate reflection of the original structure of the NFR scale.[23,25] According to COSMIN group,[23,25] the original long version can be considered the "gold standard". Internal consistency of the final structure was evaluated with Cronbach's alpha.[23] As well as in the criterion validity, it was decided *a priori* that the final structure must have good internal consistency (>0.70).Finally, all analyses were performed using the software R v4.1.1[30] using '*EFA.MRFA*' v1.1.2,[34] '*psych*' v2.1.6[35] and '*lavaan*' v0.6–9[36] packages.

Results

Characteristics of respondents

The complete sample consisted of 672 workers (528 females [78.6%] and 144 males [21.4%]) with a mean age of 30.7 ± 9.0 years and NFR index of 37.5 ± 17.5. After the split-half validation method, these variables were consistent across the development and validation samples (development sample: 259 females [77.1%], 77 males [22.9%], mean age of 30.6 ± 9.0 years, mean NFR index of 37.6 ± 17.2; validation sample: 269 females [80.1%], 67 males [19.9%], mean age of 30.8 ± 8.9 years, mean NFR index of 37.3 ± 17.8). Among the 672 participants, 11 cases with missing responses were identified, seven in the development sample and four in the validation sample. These cases were treated by means of list-wise deletion (i.e., the individual with one or more missing item was excluded). Thus, 661 participants were included in the data analysis. The distribution of the missing responses is presented in Table S1 of Supplementary Online Materials, as well as the distribution of responses for each item in Fig. S1 of Supplementary Online Materials. The correlation matrix for each item of the scale, of the complete sample, is presented in Table S2 of Supplementary Online Materials.

Exploratory factor analysis

The development sample data set met the assumptions to proceed with the EFA. The Bartlett's test of sphericity yielded significant results, $\chi^2(55)=1852.6$, $p<0.001$, and the KMO test yielded a statistic of 0.92 (ranging from 0.57 to 0.95 between items; see Table S3 of the Supplementary Online Materials). The parallel analysis (Fig. 1) that determined the number of factors to be extracted showed support for a one-factor structure.

The factor loads and communalities for the 11 items of the scale are shown in Table 2. The one-factor structure accounts for 49% of the total variance. In general, the items had factor loadings ranging from 0.61 to 0.83 and communalities ranging from 0.45 to 0.69 (all above the pre-specified criteria). The exceptions were item 4, which demonstrated low value of factor loading (0.11) and communality (0.01) and item 11, which showed a low communality value (0.38). As previously mentioned, EFA is an interactive and multistep process, therefore, the entire analysis was redone without item 4, then without item 11, and finally without both items. The same result of a construct with one-factor structure was confirmed in this analysis (data not shown). Thus, the structure without items 4 and 11 extracted by EFA was used in the following analyses.

Confirmatory factor analysis

When examining the CFA results through the goodness-of-fit indices, all structures evaluated in the study demonstrated excellent to acceptable fit on all indices (χ^2/df <5.00, CFI >0.95, TLI >0.95, and SRMR <0.08), except the RMSEA which was above the acceptable cut-off point of <0.08 (Table 3). Inspections of the standardized residual covariance matrix and the modification indices of the Brazilian structure showed that items 2 and 5 have high standardized residual covariances with several other items, suggesting they should be excluded. Furthermore, a comparison of AIC and BIC statistics among the structures revealed that the Brazilian structure without items 2, 4, 5, and 11 demonstrated superior fit compared to the other structures evaluated in the study (i.e., lower values; cf. Brazilian structure, Table 3), proving to be the most adequate structure tested.

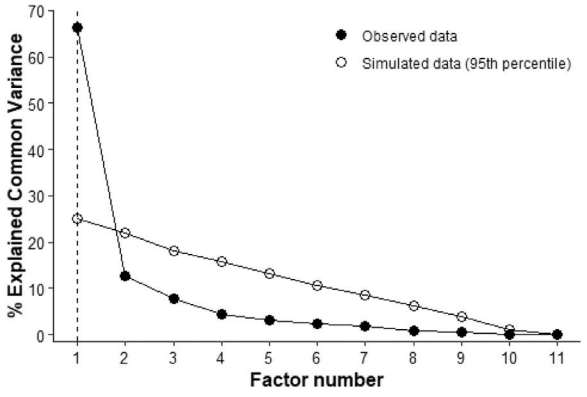

Fig. 1 Parallel analysis. The one-factor structure is shown by the vertical black dashed line. Of note — factors above the line of the simulated data (95th percentile) indicate the number of factors in the structure.

Table 2 Results of factor loads and communalities of the Exploratory Factor Analysis assuming one-factor structure using the development sample; as well as the percentage of total variance of the factor.

Items	Factor loading	Communalities
Item 1	0.73	0.53
Item 2	0.80	0.64
Item 3	0.80	0.64
Item 4	**0.11**	**0.01**
Item 5	0.72	0.52
Item 6	0.77	0.60
Item 7	0.67	0.45
Item 8	0.72	0.51
Item 9	0.67	0.45
Item 10	0.83	0.69
Item 11	0.61	**0.38**
% of variance	49	—

Note: the factor loads and communalities in bold indicate values below the pre-specified criterion.

The parameter estimates of the Brazilian structure had moderate to high standardized coefficients (i.e., >0.67) for all items. The standardized coefficients and associated data of the Brazilian structure are presented in Table S4 of Supplementary Online Materials.

Criterion validity and internal consistency

When comparing the criterion validity of the various prespecified structures (final version of the Brazilian structure with 7 items, French, Italian, Danish 1 and 2) to the original 11-item scale, our analysis showed excellent ICC values with all structures meeting our criteria for validation (ICC ≥0.95, decided *a priori*) (Table 4). Moreover, the results suggest the idea that the Brazilian structure obtained by CFA (cf. Brazilian structure, Table 3) is equivalent to the original scale of 11 items. Furthermore, the Cronbach's alpha for the Brazilian structure was 0.85, showing good internal consistency.

Discussion

The EFA confirmed the unidimensionality of the Br-NFR scale and the structure accounts for 49% of the total variance. The CFA demonstrated excellent to acceptable fit on all measured indices of the Brazilian structure without items 2, 4, 5, and 11 (i.e., Brazilian 7-item structure). The criterion validity showed that the Brazilian 7-item structure found in this study is an adequate reflection of the original scale of 11-item, and also presented excellent internal consistency. The Br-NFR with the 7 items is shown in Table S5 of Supplementary Online Materials.

Measurement properties of Br-NFR

The hypothesis of the unidimensionality of Br-NFR was confirmed by the EFA. This result is in agreement with previous studies that selected the number of factors based on the eigenvalue-over-one criterion that is less robust than the parallel analysis employed in this study.[9,11-13,16,28] On the

Table 3 Results of the adjustment indices of adequacy of the Confirmatory Factor Analysis.

Structures	χ^2	df	χ^2/df	CFI	TLI	RMSEA (90% CI)	SRMR	AIC	BIC
Brazilian	55.93	14	3.99	0.98	0.97	0.09 (0.07, 0.12)	0.04	4759.19	4812.47
Dutch	176.87	44	4.02	0.97	0.96	0.10 (0.08, 0.11)	0.05	7320.29	7404.00
French	157.15	35	4.49	0.96	0.95	0.10 (0.09, 0.12)	0.06	6781.89	6857.99
Italian	162.79	35	4.65	0.97	0.96	0.10 (0.09, 0.12)	0.06	6589.17	6665.27
Danish 1	104.44	27	3.87	0.97	0.96	0.09 (0.07, 0.11)	0.05	6287.26	6355.75
Danish 2	93.87	27	3.48	0.97	0.96	0.09 (0.07, 0.11)	0.05	6281.88	6350.37

Abbreviations: χ^2, qui-square; df, degree of freedom; χ^2/df ratio, qui-square divided by the degree of freedom; CFI, Comparative Fit Index; TLI, Tucker-Lewis Index; RMSEA, Root Mean Square Error of Approximation; CI, confidence interval; SRMR, Standardized Root Mean Square Residual; AIC, Akaike Information Criterion; BIC, Bayesian Information Criterion. Brazilian structure – one-factor structure without items 2, 4, 5 and 11; Dutch structure – original scale of 11-item (Note - structure equal to the Swedish, Iranian and Chinese scale); French structure with one-factor structure without item 10; Italian structure with one-factor structure without item 4; Danish structure 1 with one-factor structure without items 2 and 11; Danish structure 2 with one-factor structure without items 3 and 11.

other hand, this result differed to some extent from the content validity conducted in the study by Stevens and collaborators,[15] which suggested that the scale has a two-factor structure. However, the EFA performed in the study[15] showed that the scale is unidimensional, which is consistent with our results and previous literature. Furthermore, our structure explained 49% of the total variance of the scale, while the original[9] and French[11] versions explained 48% and 44% of the total variance, respectively.

The Brazilian structure demonstrated acceptable to excellent fit and lower values of AIC and BIC, compared to the other pre-specified structures evaluated in the study (original[3,9] [Dutch], Swedish,[16] Iranian,[14] Chinese,[13] French,[11] Italian,[12] Danish[15] 1 and 2), confirming our initial hypothesis that the structure extracted in this study would present acceptable goodness-of-fit indices. Thus, the Brazilian 7-item structure is a singular one compared to the literature regarding excluded items; i.e., the Italian[12] version deleted item 4, the French[11] version excluded item 10, and the Danish[15] version, merged items 2 and 3 and excluded item 11.

Regarding the criterion validity, all tested structures proved to be equivalent to the original scale of 11-item as they met our *a priori* criteria (i.e., ICC $\geq$0.95). The internal consistency of 0.85 found in this study was similar to that of other NFR (i.e., original[3,9] [Dutch], Swedish,[16] Chinese,[13] French,[11] Italian[12]) versions which ranged between 0.78 and 0.90.

Practical implications

The validation of the Br-NFR contributes to the improvement of a more valid measure, giving greater confidence in the results of future studies that use the scale, as well as for its use by occupational health services. In addition, decreasing the instrument length from 11 to 7 items reduces the response time, decreasing the burden on the respondent.[37,38] In particular, the removal of items 4 and 5 is helpful, as these items are confusing for those who work night shift and return home during the day, and for workers who work six days a week without the possibility of having two days off. Thus, the exclusion of items 2, 4, 5, and 11 is an advantage, creating greater clarity and ease of applying the Br-NFR.

Strengths and limitations

Strengths of the present study are the adequate sample size and the analysis conducted according to COSMIN consensus[25]; the implementation of the polychoric correlation matrix when using polytomous data[31,32]; and a small number of missing values in the dataset. However, some limitations need to be considered. First, the data used to calculate the criterion validity came from the same sample that responded to the 11-item scale, which could influence the results. Nevertheless, as the scale has a short length, we believe this would be unlikely to influence the results. In addition, for this study, only blue-collar workers, call center operators, and office workers were included, limiting the use of this validated structure for other occupational groups. However, this is still a very diverse sample (e.g., different shifts, working hours, working days per week), ensuring wide external validity. Notwithstanding these limitations, our study offers a contribution to research in the field of ergonomics due to the high methodological quality and the validation of the NFR structure in a large population.

Table 4 Result of criterion validity of the different pre-specified structures compared to the original 11-item scale.

	Structures				
	Brazilian	French	Italian	Danish 1	Danish 2
ICC (95% CI)	0.964 (0.956, 0.975)	0.995 (0.990, 0.997)	0.987 (0.971, 0.993)	0.990 (0.985, 0.993)	0.989 (0.978, 0.993)

Abbreviation: ICC, Intraclass Correlation Coefficient; CI, confidence interval. Brazilian structure – one-factor structure without items 2, 4, 5 and 11; French structure with one-factor structure without item 10; Italian structure with one-factor structure without item 4; Danish structure 1 with one-factor structure without items 2 and 11; Danish structure 2 with one-factor structure without items 3 and 11.

Future studies

We recommend that future studies should evaluate other measurement properties of the Br-NFR with 7 items, such as cross-cultural validity/measurement invariance, reliability, measurement error, hypothesis testing for construct validity, and responsiveness.[25] We also suggest that the structure validated in this study should be tested in other languages using CFA and that the use of a polychoric correlation matrix and coherent extraction methods for ordinal data should also be considered.[31,32]

Conclusion

The present study confirmed the unidimensionality of the Br-NFR and the structure with seven items (without items 2, 4, 5, and 11) showed excellent fit indices. Additionally, the Brazilian 7-item structure showed to be an adequate reflection of the original 11-item scale and also has good internal consistency.

Conflicts of interest

All authors declare no conflict of interest.

Acknowledgements

This research was supported by the São Paulo Research Foundation (FAPESP; grant numbers 2018/06359-4 and 2019/25140-6).

References

1. Sakowski P, Marcinkiewicz A. Health promotion and prevention in occupational health systems in Europe. *Int J Occup Med Environ Health*. 2018;32(3):353—361. https://doi.org/10.13075/ijomeh.1896.01384.
2. Boileau P-E. Sustainability and prevention in occupational health and safety. *Ind Health*. 2016;54(4):293—295. https://doi.org/10.2486/indhealth.54-293.
3. van Veldhoven M, Broersen S. Measurement quality and validity of the "need for recovery scale. *Occup Environ Med*. 2003;60 (suppl 1):i3—i9. https://doi.org/10.1136/oem.60.suppl_1.i3.
4. van Veldhoven MJPM, Sluiter JK. Work-related recovery opportunities: testing scale properties and validity in relation to health. *Int Arch Occup Environ Health*. 2009;82(9):1065—1075. https://doi.org/10.1007/s00420-009-0411-z.
5. Sluiter JK, de Croon EM, Meijman TF, Frings-Dresen MHW. Need for recovery from work related fatigue and its role in the development and prediction of subjective health complaints. *Occup Environ Med*. 2003;60:i62—i70. https://doi.org/10.1136/oem.60.suppl_1.i62.
6. de Croon EM, Sluiter JK, Frings-Dresen MHW. Need for recovery after work predicts sickness absence: a 2-year prospective cohort study in truck drivers. *J Psychosom Res*. 2003;55 (4):331—339. https://doi.org/10.1016/S0022-3999(02)00630-X.
7. de Croon EM, Sluiter JK, Frings-Dresen MHW. Psychometric properties of the Need for Recovery after work scale: test-retest reliability and sensitivity to detect change. *Occup Environ Med*. 2006;63(3):202—206. https://doi.org/10.1136/oem.2004.018275.
8. van Veldhoven M, Meijman TF. *Het Meten van Psychosociale Arbeidsbelasting Met Een Vragenlijst: De Vragenlijst Beleving En Beoordeling van De Arbeid (VBBA)*. Amsterdam: Nederlands Instituut voor Arbeidsomstandigheden (NIA); 1994. http://resolver.tudelft.nl/uuid:d231f2f3-8574-4e77-862b-4abe1ebd4df5.
9. Jansen NWH, Kant Ij, Brandt PA. Need for recovery in the working population: description and associations with fatigue and psychological distress. *Int J Behav Med*. 2002;9(4):322—340. https://doi.org/10.1207/S15327558IJBM0904_03.
10. Moriguchi CS, Alem MER, van Veldhoven M, Coury HJCG. Cultural adaptation and psychometric properties of Brazilian need for recovery scale. *Rev Saude Publica*. 2010;44(1):131—139. https://doi.org/10.1590/S0034-89102010000100014.
11. Dupret É, Bocéréan C, Feltrin M, Chemolle É, Lebon J. L'échelle de besoin de récupération: adaptation et validation françaises. *Sante Publique (Paris)*. 2018;30(4):445—454. https://doi.org/10.3917/spub.185.0445.
12. Pace F, Lo Cascio V, Civilleri A, Guzzo G, Foddai E, Van Veldhoven M. The need for recovery scale: adaptation to the Italian context. *Eur Rev Appl Psychol*. 2013;63(4):243—249. https://doi.org/10.1016/j.erap.2013.05.001.
13. Qi X, Liu Y, Ji S, Zheng R, Zhang Y, Deng H. Primary modification of Chinese need for recovery scale. *Chin J Clin Psychol*. 2014;2014(5):839—841. http://caod.oriprobe.com/articles/43193957/Primary_Modification_of_Chinese_Need_for_Recovery_Scale.htm.
14. Samadi H, Kalantari R, Mostafavi F, Zanjirani fFarahani A, Bakhshi E. Using the need for recovery scale to assess workload in mine workers and its relationship with demographics. *J Ergon*. 2017;4(4):1—7. https://doi.org/10.21859/joe-04041.
15. Stevens ML, Crowley P, Garde AH, Mortensen OS, Nygård CH, Holtermann A. Validation of a short-form version of the Danish need for recovery scale against the full scale. *Int J Environ Res Public Health*. 2019;16(13). https://doi.org/10.3390/ijerph16132334.
16. Wentz K, Gyllensten K, Archer T. Recording recovery opportunities at work and functional fatigue after work: two instruments adapted to the Swedish context. *COJ Nurs Healthc*. 2017;1 (2):1—10. https://doi.org/10.31031/COJNH.2017.01.000507.
17. Silva-Costa A, Griep RH, Fischer FM, Rotenberg L. Need for recovery from work and sleep-related complaints among nursing professionals. *Work*. 2012;41(suppl 1):3726—3731. https://doi.org/10.3233/WOR-2012-0086-3726.
18. de Negreiros AWFTRW, Rodrigues da Silva P, Arezes PMFM, Dangelino R, Padula RS. Manufacturing assembly serial and cells layouts impact on rest breaks and workers' health. *Int J Ind Ergon*. 2019;70:22—27. https://doi.org/10.1016/j.ergon.2019.01.005. November 2018.
19. Santos HG, Chiavegato LD, Valentim DP, Padula RS. Effectiveness of a progressive resistance exercise program for industrial workers during breaks on perceived fatigue control: a cluster randomized controlled trial. *BMC Public Health*. 2020;20 (1):849. https://doi.org/10.1186/s12889-020-08994-x.
20. Lidegaard M, Søgaard K, Krustrup P, Holtermann A, Korshøj M. Effects of 12 months aerobic exercise intervention on work ability, need for recovery, productivity and rating of exertion among cleaners: a worksite RCT. *Int Arch Occup Environ Health*. 2018;91 (2):225—235. https://doi.org/10.1007/s00420-017-1274-3.
21. Gupta N, Wåhlin-Jacobsen CD, Henriksen LN, Abildgaard JS, Nielsen K, Holtermann A. A participatory physical and psychosocial intervention for balancing the demands and resources among industrial workers (PIPPI): study protocol of a cluster-randomized controlled trial. *BMC Public Health*. 2015;15 (1):274. https://doi.org/10.1186/s12889-015-1621-9.
22. 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. https://doi.org/10.1016/j.jclinepi.2006.03.012.
23. de Vet HCW, Terwee C, Mokkink LB, Knol DL. *Measurement in Medicine: a Practical Guide. 1 ed*. Cambridge University Press; 2011.
24. Mokkink LB, Prinsen CAC, Bouter LM, de Vet HCW, Terwee CB. The COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) and how to select an outcome measurement instrument. *Braz J Phys Ther*. 2016;20 (2):105—113. https://doi.org/10.1590/bjpt-rbf.2014.0143.
25. Prinsen CAC, Mokkink LB, Bouter LM, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. *Qual Life Res*. 2018;27(5):1147—1157. https://doi.org/10.1007/s11136-018-1798-3.
26. Worthington RL, Whittaker TA. Scale development research: a content analysis and recommendations for best practices. *Couns Psychol*. 2006;34(6):806—838. https://doi.org/10.1177/0011000006288127.
27. Costello AB, Osborne J. Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis. *Pract Assessment, Res Eval.*. 2005;10(7):1—9. https://doi.org/10.7275/jyj1-4868.
28. Qi X, Liu Y, Zhang J, et al. Relationship between work strain, need for recovery after work and cumulative cortisol among kindergarten teachers. *Int Arch Occup Environ Health*. 2015;88 (8):1053—1059. https://doi.org/10.1007/s00420-015-1033-2.
29. Moriguchi CS, Trevizani T, de Fátima Carreira Moreira R, Januário LB, de Oliveira AB, Coury HJCG. Need for recovery assessment among nursing professionals and call center operators. *Work*. 2012;41(suppl 1):4838—4842. https://doi.org/10.3233/WOR-2012-0773-4838.
30. R. Core Team. R: a language and environment for statistical computing. Published online 2021. http://www.r-project.org/.
31. Timmerman ME, Lorenzo-Seva U. Dimensionality assessment of ordered polytomous items with parallel analysis. *Psychol Methods*. 2011;16(2):209—220. https://doi.org/10.1037/a0023353.
32. Li C-H. Confirmatory factor analysis with ordinal data: comparing robust maximum likelihood and diagonally weighted least squares. *Behav Res Methods*. 2016;48(3):936—949. https://doi.org/10.3758/s13428-015-0619-7.
33. Schreiber JB, Nora A, Stage FK, Barlow EA, King J. Reporting structural equation modeling and confirmatory factor analysis results: a review. *J Educ Res*. 2006;99(6):323—338. https://doi.org/10.3200/JOER.99.6.323-338.
34. Navarro-Gonzalez D., Lorenzo-Seva U. EFA.MRFA: dimensionality assessment using minimum rank factor analysis. Published online 2020.
35. Revelle W. psych: Procedures for psychological, psychometric, and personality research. Published online 2020.
36. Rosseel Y. lavaan: an R package for structural equation modeling. *J Stat Softw*. 2012;48(2). https://doi.org/10.18637/jss.v048.i02.
37. Sahlqvist S, Song Y, Bull F, Adams E, Preston J, Ogilvie D. Effect of questionnaire length, personalisation and reminder type on response rate to a complex postal survey: randomised controlled trial. *BMC Med Res Methodol*. 2011;11(1):62. https://doi.org/10.1186/1471-2288-11-62.
38. Rolstad S, Adler J, Rydén A. Response burden and questionnaire length: is shorter better? A review and meta-analysis. *Value Heal*. 2011;14(8):1101—1108. https://doi.org/10.1016/j.jval.2011.06.003.