

# Psychometric Properties of the Obstacles to Return-to-Work Questionnaire in a Brazilian context

Daniela Milani<sup>1,2</sup>  · Neusa Maria Costa Alexandre<sup>2</sup> ·  
Juliana Alvares Duarte Bonini Campos<sup>3</sup> · Sheilah Hogg-Johnson<sup>4</sup>

Published online: 16 November 2016  
© Springer Science+Business Media New York 2016

**Abstract** *Purpose* To evaluate the psychometric properties of the Obstacles to Return-to-Work Questionnaire (ORTWQ) among Brazilian workers on sick-leave due to musculoskeletal disorders. *Methods* Confirmatory factor analysis was conducted to evaluate the factor structure validity of the ORTWQ. Model fit indices and salience of factor loadings were assessed. The convergent validity was estimated using the Average Variance Extracted (AVE) and Composite Reliability (CR). The correlational analysis was verified using the Spearman Correlation between the ORTWQ and other specific tools. Discriminant Validity, internal consistency, stability (test–retest) and floor/ceiling effect were also assessed. *Results* A total of 301 participants completed the ORTWQ with a mean age of 45.0 (9.9) years. After refinement, the factor structure indexes of the oblique model were [ $\chi^2/df = 1.8$ ; CFI = 0.9; TLI = 0.9; PGFI = 0.7; PCFI = 0.8; RMSEA = 0.05 (90% CI 0.05–0.06)]. Only Depression, Physical Workload and Perceived Prognosis subscales presented suitable AVE indices: 0.63, 0.51 and 0.52 respectively. The correlations between ORTWQ and the other questionnaires were appropriate for almost all subscales. Reliability evaluation showed adequate estimates for all subscales except for the

Worry Due to Sick-Leave (CR = 0.45;  $\alpha = 0.44$ ; ICC = 0.69). A higher order hierarchical model is suggested, in order to estimate an overall score to ORTWQ in a Brazilian population. *Conclusions* The psychometric properties of the Brazilian version of the ORTWQ were evaluated and after refinement, the validity, reliability and floor/ceiling effects indexes were suitable when applied to a sample of Brazilian workers on sick-leave due to musculoskeletal disorders. However, the factor structure presented some issues regarding convergent and discriminant validity.

**Keywords** Return to work · Questionnaires · Reproducibility of results · Validation studies as topic · Musculoskeletal diseases

## Introduction

Musculoskeletal disorders (MSD) represent a serious public health problem in many countries, affecting workers from a range of occupations [1, 2]. They are among the most responsible for work disability, work sick-leave and high social insurance payment costs [3, 4]. Studies [2, 5–7] have shown that individual characteristics, biomechanical, psychosocial and organizational workplace factors contribute to MSD and to work disability. These factors play an important role in return-to-work (RTW) outcomes after a sick-leave and can represent obstacles to come back to previous job tasks. [2, 5–7].

Conceptual models used to frame work disability research and interventions have progressed from a strictly biomedical concept of work disability and rehabilitation to incorporate multiple facets of the worker, their work, the work environment and society as a whole [3, 8–11].

✉ Daniela Milani  
danmilani84@gmail.com

<sup>1</sup> Nursing Department at Universidade Estadual do Centro-Oeste (Unicentro), Guarapuava, PR, Brazil

<sup>2</sup> Nursing Faculty at University of Campinas (Unicamp), Campinas, SP, Brazil

<sup>3</sup> Departamento de Odontologia Social, Faculdade de Odontologia de Araraquara, Universidade Estadual Paulista (Unesp), Araraquara, SP, Brazil

<sup>4</sup> Institute for Work and Health, Toronto, ON, Canada

However the assessment and identification of obstacles to RTW continue to be poorly studied. Recommendations have been made to conduct studies that provide an international scientific evidence-base to develop appropriate tools [8] and to identify barriers and facilitators for RTW in order to implement appropriate screening measures [3]. Therefore, instruments that evaluate obstacles to RTW could provide important information for health professionals, to guide their decisions in selecting appropriate interventions and treatments for patients during a rehabilitation program [7, 12].

The “Obstacles to return-to-work questionnaire” (ORTWQ) was originally developed in Sweden and its purpose was to identify the barriers to RTW in patients with MSD in the chronic phase of disability [5]. The original version presented appropriate psychometric properties in the Swedish sample tested, and was found to be a good predictor of an episode of sick leave and has demonstrated satisfactory validity when compared to other instruments [5, 7]. The ORTWQ has 55 items across nine subscales, which are sectioned into three parts: Part I—Depression (4 questions) and Pain Intensity (4 questions); Part II—Difficulties at Work Return (8 questions), Physical Workload and Harmfulness (8 questions), Social Support at Work (6 questions); Worry Due to Sick Leave (3 questions); Work Satisfaction (9 questions); Family Situation and Support (7 questions); and, Part III—Perceived Prognosis of Work Return (6 questions). The response scale is a 7-point Likert type scale with text anchors at both ends (e.g. Not at all—very much; never—always; does not agree at all—agree completely; no chance—very big chance, etc.). Thirteen items (Part II: 2, 5, 9, 13, 18, 27, 31, 32; Part III: 1, 3, 4, 5, 6) are formulated in a negative way and their response scale should be reversed prior to analysis of the item or before calculating the subscales or overall ORTWQ scores.

The study that originally presented the ORTWQ [5] provided an English language translation of the instrument that was developed through a forward and backward translation process. In the Brazilian context the cross-cultural adaptation process of the ORTWQ followed all five steps of the recommended methodological standards [13, 14]: translation, synthesis, back translation, experts committee evaluation and pre-test. A previous study translated the Brazilian version of the ORTWQ [15]. Minor changes were needed in order to fit the Brazilian context which demonstrated adequate conceptual, semantic and idiomatic equivalences when compared to the original version [15].

Since different population’s characteristics may influence directly the way an instrument is able to assess a construct, it is crucial to evaluate the psychometric properties for that population and its desired purpose before

using an adapted tool [16–18]. Prior to this study, the psychometric properties of the ORTWQ had not been conducted in the Brazilian population. Therefore, the purpose of this study was to evaluate the psychometric properties of the ORTWQ in a Brazilian sample.

## Methods

### Settings, Subjects and Data Collection Procedures

This study used a convenience sample recruited in two publicly funded occupational health clinic settings between July 2013 and April 2014. The subjects were Brazilian workers who met the following eligibility criteria: 18 years old or over; diagnosed with a MSD; and, at recruitment, had a recorded absence from work attributed to a MSD within the previous 60 days. All who met the criteria (320 subjects) were invited by the researcher to answer a series of questionnaires while they were waiting for their appointment with the physician, physiotherapist and/or occupational therapist. All the interviews were conducted by the same researcher (DM).

### Instruments

We used the Brazilian version of the *Obstacles to Return to Work Questionnaire* as described previously [15]. In order to evaluate the concurrent validity of the nine subscales of the ORTWQ we employed established measures with similar constructs from several other instruments including: *World Health Organization Quality of Life Instrument (WHOQoL)* [19, 20], the *Numeric Rating Scale for Pain* [21, 22], the *Survey of Pain Attitudes Brief* [23, 24], the *Need for Recovery Scale (NFRS)* [25, 26], the *Fear Avoidance Beliefs Questionnaire (FABQ)* [27, 28], the short version of the *Job Stress Scale (JSS)* [29, 30], and the *Occupational Stress Indicator (OSI)* [30, 31]. These scales are briefly described below. In order to be chosen for this study, all of these scales have been properly cross-culturally adapted for use in Brazil and their psychometric properties have been previously tested in at least one Brazilian sample.

The *World Health Organization Quality of Life Instrument (WHOQoL)* [19, 20] is a generic quality of life instrument which consists of six domains constructed from 24 subscales and 100 items. For this study we used three subscales: Pain (4 items), Negative Feelings (4 items) and Social Support (4 items). The *Numeric Rating Scale for Pain* [21, 22] is constructed by averaging responses of two items related to pain in the last 3 months and pain in the last seven days with item scores ranging from zero (no pain

at all) to ten (worst possible pain). *The Survey of Pain Attitudes Brief (SOPA)* [23, 24] is a 30-item instrument that contains seven domains about pain attitudes. For our study we used the Disability (3 items) and Medical Cure (5 items) domains. The *Need for Recovery Scale (NFRS)* [25, 26] is a unidimensional scale composed of 11 questions. An evaluation of NFRS among workers revealed that it is sensitive to quantifying occupational workloads for an individual such as mental and physical work demands, lack of decision latitude and number of working hours [26]. The *Fear Avoidance Beliefs Questionnaire (FABQ)* [27, 28] assesses individuals' beliefs and fear of pain via two subscales related to their physical activities (5 items) and work (11 items). For our study, only the work-FABQ subscale was used. The *Job Stress Scale (JSS)* short version [29, 30] measures demands, control and support at work. The entire tool has 17 questions combined into three subscales: Demands (5 items), Control (6 items) and Support (6 items). For this study we only used the Support subscale. The *Occupational Stress Indicator (OSI)* [31, 32] consists of 171-items across six domains: Sources of Pressure in the Job (61 items), Type A Behavior (14 items), Coping Skills (28 items), Locus of Control (12 items), Job Satisfaction (22 items) and Health [Mental Health (18 items) and Physical Health (16 items)]. In this study we used only the Job Satisfaction subscale. More detail on how the ORTWQ constructs were compared with these instruments are described in the concurrent validity session.

Data on sociodemographic and work characteristics of the sample were also collected. A body diagram from the *Nordic Musculoskeletal Symptoms Questionnaire* [33] was administered so that participants could indicate specific bodily locations for any pain they experienced in the: (1) last 12 months and (2) last 7 days.

### Sample Size

The minimum sample size was estimated based on the recommendations of Hair et al. [34] and Kim [35] who suggest between 5 and 10 subjects per item of the instrument are required to conduct confirmatory factor analysis. Given that ORTWQ has 55 items, the sample should have at least 275 respondents. For concurrent validity analysis, the sample was estimated by the coefficient of correlation method for sample size calculation [36] and a total of at least 175 subjects were required to fill the other specific instruments besides the ORTWQ. In order to evaluate stability (test–retest), the sample size estimate followed the recommendations of Terwee et al. [37]. They suggest a sample of 50 subjects as being sufficient to assess test–retest reliability.

Considering all the requirements described above, we invited 320 subjects to participate in this study. We had a total of 301 responses (response rate 94.3%) to the

ORTWQ. Of this total, 185 were invited to complete the other instruments mentioned previously. In order to check the stability (test–rest), 51 participants, who completed the first round, were also invited to complete the ORTWQ a second time.

### Psychometric Analyses

#### *Psychometric Sensitivity and Generalizability*

Each ORTWQ item was evaluated via summary descriptive statistics (mean, median, mode, and standard deviation) and form (skewness and kurtosis). The psychometric sensitivity was considered adequate when the absolute value of skewness was less than 3 and kurtosis was less than 7 [38, 39]. Tests on generalizability were verified in the subsamples in order to check if they were similar to the total sample.

#### *Construct Validity*

Construct validity was assessed through factor structure validity, convergent validity and discriminant validity.

A Confirmatory Factor Analysis (CFA) by the Maximum Likelihood parameter estimation method was conducted in order to verify the degree to which the Brazilian version of the ORTWQ satisfied the same orthogonal factor structure as the original questionnaire (55 items combined into nine subscales organized in three parts) [5]. Literature has discussed the RTW and the Obstacles to RTW as being a multidimensional theoretical model [8, 12, 40]. It is known that an instrument that aims to evaluate a multidimensional theoretical model presents subscales or parts with some degree of relationship [41]. Oblique models (as result of an oblique rotation) are often seen as producing more appropriate solutions for research in social science, human behavior, etc. [42]. Thus, in order to verify if an oblique rotation would produce a better solution than the *orthogonal first order model* [5], we also tested an *oblique model*.

This study evaluated the adequacy of the models by following the Goodness-of-fit indices: Chi square by degrees of freedom ratio ( $\chi^2/df$ ), *Comparative fit index* (CFI), *Tucker-Lewis Index* (TLI), *Parsimony Goodness of Fit Index* (PGFI), *Parsimony Comparative fit index* (PCFI), and *Root Mean Square Error of Approximation* (RMSEA) with confidence intervals of 90% [90% CI]. These parameters were considered adequate when  $\chi^2/df \leq 2.0$ ;  $CFI \geq 0.9$ ,  $TLI \geq 0.9$ ,  $PGFI$  and  $PCFI \geq 0.6$ ,  $RMSEA$  [90% CI]  $< 0.1$  [34, 38]. In addition, factor loadings were assessed for saliency [Standardized estimates ( $\lambda \geq 0.4$ )]

and statistical significance ( $p < 0.05$  expected). We removed items with  $\lambda < 0.4$ . Model modifications were considered based on the modification indices that were estimated by the Lagrange Multipliers Method as described by Maroco [38].

The convergent validity was estimated by the Average Variance Extracted (AVE). Values of  $AVE \geq 0.5$  indicate adequate convergent validity [34, 38, 43]. The discriminant validity assesses whether the items that reflect a factor are not strongly correlated with another factor [38, 43]. It was computed by correlational analysis, and considered adequate if  $AVE_i$  and  $AVE_j \geq \rho_{ij}^2$  [38, 43].

#### Concurrent Validity

The concurrent validity is calculated in order to evaluate the theoretical approximation between factors/instruments. It was estimated by correlational analysis between ORTWQ subscales and other measurement instruments that present similar constructs, namely: ORTWQ Pain Intensity with Numeric Rating Scale for Pain and WHO-QoL Pain subscale; ORTWQ Difficulties at Work Return with NFRS and the SOPA Disability Scale; ORTWQ Physical Workload and Harmfulness with FABQ-Work scale; ORTWQ Social Support at Work with the JSS Support Scale; ORTWQ Work Satisfaction with OSI-Job Satisfaction scale; ORTWQ Family Situation and Support with WHOQoL Social Support subscale; and the ORTWQ Perceived Prognosis of Work Return with items 15 and 16 of the FABQ and with SOPA-Medical Cure subscale. For this analysis we used Spearman Correlation analysis ( $r$ ) and for each paired construct, we expected at least a moderate magnitude correlation ( $r > 0.3$ ) but preferably a strong correlation ( $r > 0.5$ ) [44] as an indicator of concurrent validity.

#### Reliability

Reliability was assessed by internal consistency, Composite Reliability (CR) and stability (test–retest). The internal consistency was assessed using the standardized Cronbach's alpha coefficient ( $\alpha$ ). The Composite Reliability (CR) was estimated according to Fornell and Larcker [43]. These criteria were considered adequate when  $\alpha \geq 0.7$  [35] and  $CR \geq 0.7$  [34, 37, 42]. The ORTWQ was administered twice, over an interval of between seven to fifteen days (mean time: 11.6), with no anticipated change in the construct to be measured between both occasions [36]. This criteria was assessed using the Intraclass Correlation Coefficient (ICC) for each subscale separately and was considered adequate when  $ICC \geq 0.7$  [37].

#### Floor/Ceiling Effects

Floor and Ceiling effects indicate the extent to which a score is located at the bottom or top of the scale range. We adopted the commonly used 15% threshold to identify the percentage of the sample that have the lowest and the highest scores of overall ORTWQ and its subscales to define a ceiling and floor effect [37].

#### The Higher Order Hierarchical Models and the Global Scores

After fitting of both of the models, the *orthogonal first order* and the *oblique*, the hierarchical second-order (SOHM) and third-order (TOHM) models were fitted respectively [38] aiming to obtain a global score on the ORTWQ as originally proposed by Marhold and colleagues [5], but preserving the theoretical differentiation of items. We also aimed to obtain a global score for each subscale (e.g. Depression, Pain, Difficulties to Return to Work, Social Support at Work, etc.).

The global scores of the factors (subscales) and of the ORTWQ were estimated for the models using a regression weights matrix of the model obtained in the confirmatory factor analysis [38]. Each item weight of the ORTWQ, in its different subscale structures, should be multiplied by the response given to the item, and all items must be summed to generate an overall score for the individual subscales or to obstacles to RTW.

For all statistical tests, the significance level was 5%. Statistical analyses were performed using IBM SPSS Statistics (v.20, SPSS An IBM Company, Chicago, IL) and AMOS 20.0 (SPSS Inc., Chicago, IL).

#### Ethical Aspects

This research project was submitted and approved by the University Ethics Committee. Participants signed two copies of the consent form. The ORTWQ developers [5] authorized the Brazilian cross-cultural adaptation and its future use in Brazil.

## Results

#### Participants

The mean age (SD) of the 301 participants (response rate 94.3%) was 45.0 (9.9) years. They performed several occupations with various job demands including: 48.8% physically demanding (e.g. metallurgical, general

**Table 1** Sociodemographic, work-related and health characteristics of study participants (n = 301)

Attribute	n Available	Mean (SD) or %
<i>Sociodemographic and life style characteristics</i>		
Age	301	45.0 (9.9)
Gender	301	
Male		42.5
Female		57.5
Education (in years)	301	
≤8 years		21.9
9–11 years		50.5
>11 years		27.6
Marital status	301	
Married or living with partner		59.1
Divorced or separated		20.6
Widowed		2.3
Single		17.9
<i>Work-related variables</i>		
Time in the same occupation	301	14.4 (10.2)
Work demand	301	
Physical		48.8
Mixed		25.9
Mental		25.5
Type of contract	301	
Public company		41.9
Private company		58.1
Time on sick leave (in days)	301	122.5 (469.7)
≤15 days		63.1
>15–90 days		18.6
> 90 days		18.3
Previous work sick-leave <sup>a</sup>	246	
Yes		50.4
No		49.6
Time of previous work sick-leave <sup>b</sup>	121	424.1 (631.0)
≤15 days		25.6
>15–90 days		18.2
>90 days		55.4
<i>Health-related variable</i>		
Site of disorders responsible for the sick-leave	301	
Upper limb		39.5
Back		28.9
Shoulder + back (associated)		8.3
Others		23.3

<sup>a</sup> 55 Subjects did not want to say or did not know to say if they were on sick-leave previously

<sup>b</sup> 3 Participants did not know to report how many days they were in sick leave previously

services), 25.5% mentally demanding (e.g. administrative area) and 25.9% mixed work demands (e.g. health sector). Table 1 presents a description of participant characteristics.

The mean (SD) time to complete the ORTWQ interview was 13 (4.1) min (Range = 6–26 min).

### Generalizability and Psychometric Sensitivity

There were no statistical differences between the total sample (n = 301) and the subsamples (n = 185) and/or (n = 51) characteristics.

Descriptive characteristics of the ORTWQ items, (means, standard deviation, kurtosis, skewness and frequency distribution) are presented in Table 2. There was no missing data, all response options were endorsed (min = 0 max = 6) showing distribution across the potential range of responses. There was not a high frequency (90%) of responses at floor or ceiling for any item. Skewness and kurtosis for all items met the distributional requirements to conduct the CFA using maximum likelihood.

### Construct Validity

We observed that the orthogonal first order model (*non-modified*) did not have adequate fit indices for the Brazilian sample and refinements were required. The results of the Confirmatory Factor Analysis of the *non-modified* model are presented in Table 3 and its structure is depicted in Fig. 1a.

In Part I, the DEP5 item had a higher loading on the Pain Intensity subscale than the Depression subscale. Therefore, this item was moved to the Pain Intensity subscale, since the theoretical underpinning of it relates to pain. According to modification indices, and as supported by theory, one pair of error correlations was also needed between the non-random errors of the PAIN2 and PAIN3 items.

In Part II, the items of the Family Situation and Support subscale needed refinement as the items, presented different theoretical concepts/constructs. First, their content was different and second, their loadings for the model were either high (four items) or very low (three items). Rather than remove these items, we split them across two different subscales—one for Family Support and one for Family Situation. This revised model was then fitted to the data. This led to some improvement in fit for Part II, but overall fit was still assessed as inadequate ( $\chi^2/df = 2.3$ ; CFI = 0.8; TLI = 0.8; PGFI = 0.7; PCFI = 0.7; RMSEA = 0.07 [90% CI 0.06–0.07]). In addition, the following items presented factor loadings less than 0.4:  $\lambda_{SSW31}=0.19$ ;  $\lambda_{WDSL7} = 0.37$ ;  $\lambda_{WS13} = 0.17$ ;  $\lambda_{WS36} = 0.27$ ;  $\lambda_{WS39} = 0.39$ ;  $\lambda_{FAM14} = 0.09$ ;  $\lambda_{FAM25} = 0.11$ ;  $\lambda_{FAM41} = 0.06$ .

For the *refined model* (presented in Table 3) of Part II, first five items were dropped: WS13, SSW31, WS36, FAM38 and WS39. The item WDSL7 from the Worry Due to Sick-Leave subscale also had a loading  $\lambda < 0.4$  and was a candidate for exclusion. However, we retained this item because the WDSL subscale only has three items and the loadings for the remaining two items decreased when the WDSL7 item was removed. Then, since the Family Support subscale presented no significant correlations between other subscales, it was a candidate for exclusion. Furthermore, the WS15 and PWH28 items presented very high correlations with more than two subscales (i.e. high cross-loadings), thus, they were also dropped from the model,

even though their factor loadings were  $\lambda > 0.4$ . The refined model of Part II also included error correlations for three pairs of non-random errors as suggested by modification indices (MI)  $\geq 11$ . Refinement of Part III included the removal of the PPRTW6 item ( $\lambda_{PPRTW6} = 0.19$ ) and an additional pair of error correlations between the non-random errors of the PPRTW1 and PPRTW4 items.

The results of the confirmatory factor analysis for the *Oblique model refined* are presented in Table 3 and its structure is depicted in Fig. 1c. After applying refinements, as was done for the *Orthogonal model*, the *Oblique model* showed adequate fit indices.

Convergent Validity results (presented in Table 3) found that the majority of the subscales presented AVE values below the recommended cut-off in any tested model (orthogonal or oblique). The complete 55-item model (*non-modified*) presented recommended values of AVE for Depression, Physical Workload and Harmfulness subscales. For the refined models, besides Depression, Physical Workload and Harmfulness subscales, the Perceived Prognosis of Return to Work subscale achieved an AVE  $\geq 0.5$ .

Since the results of AVE<sub>i</sub> and AVE<sub>j</sub> were  $\geq (\rho_{ij}^2)$ , the discriminant validity in this sample was not demonstrated for the Pain, Difficulties to Return-to-Work and Social Support at Work subscales.

### Concurrent Validity

The correlational analysis between the ORTWQ subscales of the refined Brazilian models and the other questionnaires is presented in Table 4. There was no scale available related to the construct of Worry Due to Sick Leave for this analysis. All except one of the ORTWQ subscales presented correlations of strong magnitude with the questionnaires that evaluated similar constructs, as hypothesized. The exception was the Family Situation subscale in the refined models. Correlation between the WHOQoL-Social Support and the Family Situation was weak ( $r = -0.2$ ).

### Reliability

Internal consistency and stability (test–retest) results are displayed in Table 3. The mean time between test and retest was 11.6 days. Cronbach's alpha coefficient and CR values were adequate ( $\geq 0.7$ ) for the majority of the subscales in all presented models. Exceptions for the *non-modified model* included the subscales Worry Due to Sick Leave ( $\alpha = 0.44$  and CR = 0.45) and Family Situation and Support ( $\alpha = 0.61$  and CR = 0.53). In the *refined models* only the Worry Due to Sick Leave still presented inadequate values.

**Table 2** Descriptive characteristics of the ORTWQ items: mean, standard deviation (SD), kurtosis (Ku), skewness (Sk) and frequency distribution of the items (n = 301)

Scale	Item <sup>a</sup>	Mean	SD	Ku	Sk	Frequency distribution (%)						
						0	1	2	3	4	5	6
Depression (DEP)	5. How often do you think about your pain?	4.4	1.8	0.0	-1.0	6.0	3.3	4.3	15.6	13.3	14.6	42.9
	6. How depressed have you felt during <i>the last week</i> ?	4.1	2.1	-0.6	-0.8	12.6	3.7	4.3	12.6	13.3	15.3	38.2
	7. How depressed have you felt, on average, during <i>the last three months</i> ?	3.9	1.9	-0.5	-0.7	8.31	4.3	7.3	17.9	16.3	17.6	28.2
	8. How often do you feel that it is difficult to start activities?	3.6	2.1	-1.1	-0.5	14.3	6.6	8.3	14.3	12.0	17.3	27.2
Pain intensity (PAIN)	1. How much pain have you had during <i>the last week</i> ?	4.6	1.4	-0.3	-0.8	0.3	2.3	6.3	12.3	21.6	16.3	40.9
	2. How much pain have you had, on average, during <i>the last three months</i> ?	4.7	1.4	0.6	-1.0	1.0	1.7	3.7	12.0	20.6	20.6	40.5
	3. How often have you been in pain during <i>the last three months</i> ?	4.8	1.5	0.1	-1.0	0.7	2.0	6.0	12.6	14.0	14.0	50.8
	4. How much does your pain prevent you from engaging in activities in your daily life?	4.5	1.6	0.4	-1.0	4.0	2.3	5.3	13.3	19.9	17.3	37.9
Difficulties at work return (DWR)	3. On return to work, fatigue will be a problem for me	3.6	2.3	-1.3	-0.5	21.6	4.7	5.0	10.3	11.3	17.3	29.9
	8. Increased ache will be a problem for me on return to work	4.5	2.0	0.1	-1.2	9.0	2.3	5.7	8.6	11.3	10.6	52.5
	11. After work I won't have the energy to do anything but rest	3.6	2.4	-1.4	-0.4	21.6	5.3	6.0	11.0	7.3	12.0	36.9
	16. I have spent time with less friends and acquaintances since I went on sick leave	2.8	2.7	-1.8	0.1	42.5	2.3	2.3	8.3	6.0	7.0	31.6
Physical workload harmfulness (PWH)	20. I have had to give up leisure activities because of my pain	4.4	2.3	-0.5	-1.1	17.3	1.0	4.0	5.0	5.7	9.3	57.8
	21. I am too sick/disabled to work professionally	2.8	2.5	-1.7	0.0	37.9	2.7	5.0	9.3	9.0	9.3	26.9
	22. Life feels meaningless when I don't have the energy to do anything after work	3.6	2.5	-1.4	-0.5	24.9	3.7	4.3	6.3	11.0	10.3	39.5
	23. I have to rest when my pain increases	5.1	1.6	3.2	-2.0	4.3	0.7	2.0	8.3	7.6	9.0	68.1
Social support at work (SSW)	1. My work demands physical effort	4.5	2.0	-0.3	-1.1	7.0	4.3	8.3	9.6	5.3	10.6	54.8
	4. I have too much to do at work	4.3	1.8	-0.1	-0.9	6.6	3.0	6.0	15.0	14.6	16.6	38.2
	12. I won't be able to reduce my sick leave since my work demands so much physical effort	3.0	2.5	-1.6	-0.1	34.2	2.0	5.3	11.3	9.6	9.6	27.9
	17. My work is the cause of my pain.	4.0	2.4	-1.2	-0.7	21.3	2.3	4.3	7.6	6.6	8.6	48.8
Worry due to sick-leave (WDSL)	19. The repetitive movements (for example with arms and hands) that my work contains aggravate my pain	4.7	2.1	0.6	-1.4	12.0	0.7	3.0	6.6	7.0	7.0	63.8
	28. My work is detrimental to my health	3.0	2.5	-1.7	-0.0	33.6	4.3	5.3	9.0	10.6	5.7	31.6
	33. If I had had another kind of job I would never have gotten any pain.	2.8	2.6	-1.7	0.1	39.9	3.0	4.7	9.3	5.7	5.0	32.7
	34. One day at my job contains many heavy work tasks	3.4	2.4	-1.5	-0.3	23.9	6.3	6.6	9.3	8.3	10.0	35.6
Worry due to sick-leave (WDSL)	5. My job supervisor has understanding for my pain problem <sup>b</sup>	2.8	2.5	-1.6	0.2	32.6	9.0	9.6	9.6	7.0	3.3	28.9
	6. There are often conflicts at my workplace	3.5	2.3	-1.4	-0.3	20.3	5.3	9.0	9.3	11.0	11.3	30.6
	18. My job supervisor tries to support me and make things easier for me at the workplace <sup>b</sup>	3.2	2.5	-1.6	-0.1	24.6	10.6	7.6	11.3	6.0	4.7	35.2
	24. My workplace has a tense atmosphere	3.9	2.1	-1.2	-0.7	18.9	3.7	4.3	11.3	8.3	7.6	45.8
Worry due to sick-leave (WDSL)	31. I get along well with my work colleagues <sup>b</sup>	0.9	1.5	3.6	2.1	64.1	16.0	6.3	6.3	1.7	1.3	4.32
	40. It feels bad that my work colleagues don't understand my pain	3.3	2.6	-1.7	-0.2	31.6	5.0	2.3	7.3	8.6	10.6	34.6
	7. I worry about what my work colleagues think about my sick leave	3.2	2.6	-1.7	-0.2	32.6	3.0	5.0	6.0	10.0	10.0	33.6
	29. It feels bad that I get so far behind at work when I'm on sick leave	3.1	2.5	-1.7	-0.2	33.2	2.7	5.0	7.6	10.3	9.6	31.6

**Table 2** continued

Scale	Item <sup>a</sup>	Mean	SD	Ku	Sk	Frequency distribution (%)							
						0	1	2	3	4	5	6	
Work satisfaction (WS)	37. I have to work hard when I get back to make up for everything I missed during sick leave	2.1	2.5	-1.3	0.6	49.5	5.3	6.6	8.6	3.3	7.0	19.6	
	2. I really enjoy my work <sup>b</sup>	1.2	1.8	1.1	1.5	57.5	11.6	9.6	10.0	2.7	1.7	7.0	
	10. My work tasks are boring	2.1	2.3	-1.2	0.5	43.9	5.7	9.6	12.0	6.6	6.6	15.6	
	13. My job is varied <sup>b</sup>	2.7	2.5	-1.6	0.2	33.6	9.0	8.30	12.3	5.0	3.0	28.9	
	15. I would like to change work tasks	3.1	2.6	-1.8	-0.1	35.2	4.0	7.0	5.0	7.0	7.3	34.6	
	26. What I do outside the job hours feels more important than my work	2.3	2.3	-1.3	0.4	41.2	4.0	7.7	18.6	7.0	5.0	16.6	
	30. I feel bored by my work	2.2	2.4	-1.4	0.5	45.2	6.3	5.3	11.0	7.6	7.3	17.3	
	35. The drawback to working is that I don't have time for other things	2.6	2.4	-1.5	0.2	36.5	5.0	5.3	14.3	11.6	8.3	18.9	
	36. I find my work schedule unsatisfactory.	2.4	2.5	-1.6	0.4	47.5	1.3	4.7	11.0	5.7	8.6	21.3	
	39. When I work I have too little time for my family/friends	3.6	2.3	-1.3	-0.5	21.6	3.0	5.7	12.3	13.0	9.6	34.9	
Family situation and support (FAM)	9. I can stand my pain thanks to the support of my friends/family <sup>b</sup>	2.3	2.3	-1.2	0.6	34.2	14.3	10.0	14.6	3.3	2.7	20.9	
	14. It's hard to find the energy to work since my family situation/social situation is bad	1.4	2.1	-0.1	1.2	63.1	5.3	4.7	7.3	5.3	4.7	9.6	
	25. Since my children take up so much energy, it's hard to find the strength to work	1.2	1.9	0.7	1.4	62.5	6.6	9.6	6.7	3.7	5.0	6.0	
	27. I feel that my family/friends have understanding for my situation <sup>b</sup>	1.4	1.9	0.7	1.3	52.8	13.0	12.0	8.3	4.0	1.0	9.0	
	32. When I am in pain, my family/friends think of something to do to make me forget my pain <sup>b</sup>	2.2	2.3	-1.1	0.7	36.2	15.6	10.6	11.3	4.3	3.0	18.9	
	38. When I am in pain, my family/friends can't stand being around me	1.5	2.2	-0.5	1.1	60.6	5.3	5.3	6.6	5.3	4.7	12.3	
	41. I need to stay home because of my family situation	0.9	1.7	2.5	2.0	75.4	5.0	3.3	5.3	2.3	3.3	5.3	
	Perceived prognosis of return-to-work (PPRTW)	1. How much chance do you think there is that you could return to work? <sup>b</sup>	1.6	2.2	-0.3	1.1	55.2	9.6	8.6	6.3	2.7	2.3	15.3
		2. How do think it's going to be to return to work?	3.3	2.4	-1.5	-0.2	26.9	4.7	5.7	9.6	14.3	10.0	29.0
		3. How convinced are you that you will recover? <sup>b</sup>	1.9	2.2	-0.8	0.8	45.2	11.3	10.0	10.0	4.0	4.3	15.3
4. How much chance is there that you will be able to return to work within 6 months? <sup>b</sup>		1.5	2.1	0.1	1.2	56.5	9.6	9.0	7.6	3.0	0.7	13.6	
5. How do you feel when you think about your possibilities for returning to work? <sup>b</sup>		1.8	2.3	-0.8	0.9	48.8	10.6	9.0	9.0	2.3	2.7	17.6	
6. How do you feel when you think about your possibilities to manage without medical assistance in the future? <sup>b</sup>		4.7	2.2	0.0	-1.3	11.3	4.3	4.0	6.3	2.0	2.7	69.4	

<sup>a</sup> There were no missing data for any item

<sup>b</sup> Items that need to be reversed in order to calculate the scores have been reversed for display in this table



**Table 3** Model fit statistics, average variance extracted, internal consistency, test–retest and floor/ceiling effects of ORTWQ's Brazilian version: orthogonal and oblique models

Models	CFA $\lambda$	$\chi^2/df$	CFI	TLI	PGFI	PCFI	RMSEA (90% CI)	Reliability			Floor/ Ceiling Effect		
								AVE	CR	$\alpha$	ICC	FE n (%)	CE n (%)
<i>Orthogonal 1st order (non-modified)</i>													
Part I	<b>0.52–0.86</b>	<b>6.44</b>	<b>0.90</b>	<b>0.85</b>	<b>0.48</b>	<b>0.61</b>	<b>0.14 (0.11–0.16)</b>				0 (0)	0 (0)	
Depression	0.52–0.86							0.53	0.82	0.80	0.74	34 (11.3)	0 (0)
Pain intensity	0.65–0.74							0.47	0.78	0.78	0.80	53 (17.6)	0 (0)
Part II	<b>0.01–0.79</b>	<b>2.55</b>	<b>0.74</b>	<b>0.72</b>	<b>0.65</b>	<b>0.69</b>	<b>0.07 (0.07–0.08)</b>						
Difficulties to return-to-work	0.55–0.77							0.40	0.84	0.84	0.86	21 (7.0)	2 (0.7)
Physical workload harmfulness	0.59–0.78							0.51	0.89	0.89	0.90	21 (7.0)	3 (1.0)
Social support at work	0.22–0.69							0.31	0.71	0.72	0.90	4 (1.3)	9 (3.0)
Worry due to sick leave	0.37–0.58							0.22	0.45	0.44	0.69	19 (6.3)	33 (12.0)
Work satisfaction	0.17–0.79							0.31	0.78	0.76	0.85	2 (0.7)	2 (0.7)
Family Support and Situation	0.01–0.72							0.19	0.53	0.61	0.73	0 (0)	25 (8.3)
Part III	<b>0.20–0.81</b>	<b>3.52</b>	<b>0.96</b>	<b>0.94</b>	<b>0.41</b>	<b>0.58</b>	<b>0.09 (0.06–0.13)</b>						
Perceived prognosis return-to-work	0.20–0.81							0.46	0.82	0.81	0.80	13 (4.3)	9 (3.0)
<i>Orthogonal 1st order-(Refined)</i>													
Part I	<b>0.58–0.88</b>	<b>3.42</b>	<b>0.96</b>	<b>0.93</b>	<b>0.48</b>	<b>0.62</b>	<b>0.09 (0.07–0.12)</b>						
Depression	0.70–0.88							0.64	0.84	0.83	0.74	41 (13.6)	13 (4.3)
Pain intensity	0.58–0.78							0.44	0.79	0.80	0.80	42 (14.0)	0 (0)
Part II	<b>0.38–0.83</b>	<b>1.88</b>	<b>0.90</b>	<b>0.88</b>	<b>0.72</b>	<b>0.80</b>	<b>0.05 (0.05–0.06)</b>						
Difficulties to return-to-work	0.55–0.77							0.41	0.84	0.84	0.86	21 (7.0)	4 (1.3)
Physical workload harmfulness	0.58–0.82							0.51	0.88	0.87	0.90	27 (9.0)	3 (1.0)
Social support at work	0.43–0.74							0.34	0.72	0.74	0.89	27 (9.0)	9 (3.0)
Worry due to sick leave	0.38–0.59							0.22	0.45	0.44	0.69	19 (6.3)	33 (12.0)
Work satisfaction	0.45–0.83							0.42	0.78	0.76	0.87	3 (1.0)	37 (12.3)
Family situation	0.65–0.66							0.45	0.71	0.71	0.78	6 (2.0)	138 (45.8)
Part III	<b>0.66–0.81</b>	<b>1.88</b>	<b>0.99</b>	<b>0.99</b>	<b>0.26</b>	<b>0.40</b>	<b>0.05 (0.00–0.11)</b>						
Perceived prognosis return-to-work	0.66–0.81							0.52	0.84	0.85	0.83	14 (4.7)	45 (15.0)
<i>Orthogonal 2nd order – (refined)</i>													
Part I	<b>0.58–0.88</b>	<b>3.41</b>	<b>0.96</b>	<b>0.93</b>	<b>0.48</b>	<b>0.62</b>	<b>0.09 (0.01–0.12)</b>						
Depression	0.70–0.88							0.64	0.84	0.83	0.74	41 (13.6)	13 (4.3)
Pain intensity	0.58–0.78							0.44	0.79	0.80	0.80	42 (14.0)	0 (0)
Part II	<b>0.37–0.82</b>	<b>1.98</b>	<b>0.88</b>	<b>0.87</b>	<b>0.72</b>	<b>0.81</b>	<b>0.06 (0.05–0.06)</b>						
Difficulties to return-to-work	0.56–0.76							0.41	0.84	0.84	0.86	21 (7.0)	4 (1.3)
Physical workload harmfulness	0.58–0.82							0.51	0.88	0.87	0.90	27 (9.0)	3 (1.0)

**Table 3** continued

Models	CFA $\lambda$	$\chi^2/df$	CFI	TLI	PGFI	PCFI	RMSEA (90% CI)	Reliability			Floor/ Ceiling Effect		
								AVE	CR	$\alpha$	ICC	FE n (%)	CE n (%)
<b>ORTWQ Brazilian version</b>													
Social support at work	0.40–0.73							0.34	0.72	0.74	0.89	27 (9.0)	9 (3.0)
Worry due to sick leave	0.37–0.58							0.22	0.45	0.44	0.69	19 (6.3)	33 (12.0)
Work satisfaction	0.46–0.81							0.42	0.78	0.76	0.87	3 (1.0)	37 (12.3)
Family situation	0.65–0.70							0.45	0.71	0.71	0.78	6 (2.0)	138 (45.8)
<i>Oblisque</i> —(refined)		<b>1.75</b>	<b>0.89</b>	<b>0.88</b>	<b>0.72</b>	<b>0.83</b>	<b>0.05 (0.05–0.06)</b>					0 (0)	0 (0)
<b>Part I</b>													
Depression	0.73–0.86							0.63	0.84	0.83	0.74	0 (0)	0 (0)
Pain intensity	0.61–0.79							0.43	0.79	0.80	0.80	0 (0)	0 (0)
<b>Part II</b>													
Difficulties to return-to-work	0.53–0.78							0.41	0.84	0.84	0.86	0 (0)	0 (0)
Physical workload harmfulness	0.58–0.82							0.51	0.88	0.87	0.90	0 (0)	0 (0)
Social support at work	0.39–0.73							0.35	0.71	0.74	0.89	0 (0)	0 (0)
Worry due to sick leave	0.37–0.58							0.22	0.45	0.44	0.69	0 (0)	0 (0)
Work satisfaction	0.45–0.82							0.42	0.77	0.76	0.87	0 (0)	0 (0)
Family Situation	0.65–0.70							0.45	0.71	0.71	0.78	0 (0)	0 (0)
<b>Part III</b>													
Perceived prognosis return-to-work	0.64–0.77							0.52	0.84	0.85	0.83	0 (0)	0 (0)

CFA confirmatory Factor Analysis;  $\chi^2/df$  Chi square by degrees of freedom, CFI comparative fit index, TLI Tucker-Lewis index, PGFI parsimony goodness of fit index, PCFI parsimony comparative fit index, RMSEA root mean square error of approximation, CI confidence intervals, AVE average variance extracted, CR composite reliability,  $\alpha$  Cronbach's alpha, ICC intraclass correlation coefficient, FE floor effect, CE ceiling effect

The test–retest stability evaluation was found to be adequate ( $ICC \geq 0.7$ ) for all subscales of all presented models (*non-modified and refined, orthogonal or oblique*), except for Worry Due to Sick Leave subscale that presented in the limit of adequacy ( $ICC = 0.69$ ).

### The Higher Order Hierarchical Models and the Global Scores

The *Orthogonal second-order model* (depicted in Fig. 1b) demonstrated similar fit to the *Orthogonal first-order model*. The *Oblique third-order model* (depicted in Fig. 1d) has also presented suitable fit indices [ $\chi^2/df = 1.85$ ; CFI = 0.9; TLI = 0.9; PGFI = 0.7; PCFI = 0.8; RMSEA = 0.05 (90% CI= 0.05–0.06)].

Higher order hierarchical models allow estimating a score for each questionnaire's subscale, as well as an overall score of the Obstacles to RTW construct from an algorithm. This algorithm considers the weight (W) of each item and is displayed in Table 5.

In order to generate a score for each subscale and/or an overall score for ORTWQ, the weights should be multiplied by the answers' values obtained for each of the items and then be summed together.

### Floor and Ceiling effects

The floor and ceiling effects were verified in all of the models (*non-modified and refined, orthogonal and oblique*) and the results are displayed in Table 3. The non-modified model presented a floor effect (17.6%) for the Pain Intensity subscale. In the refined orthogonal models, two subscales had floor/ceiling effects. These included the Family Situation subscale which had a substantial ceiling effect (45.8%) and the Perceived Prognosis to Return-to-Work Subscales which had a minimum ceiling effect (15%). All of the other subscales have no floor or ceiling effects. For the estimates of the total score and scores of the subscales of the *oblique models*, we considered the algorithms displayed in Table 5. Once we instituted these algorithms, no floor or ceiling effects for any subscale were present.

### Discussion

The original (non-modified) nine factor model of the Brazilian version of the ORTWQ did not adequately fit the data gathered for the Brazilian population sampled in this study. After refinement of, both the orthogonal and the oblique alternatives models, satisfactory psychometric properties for this sample were observed.

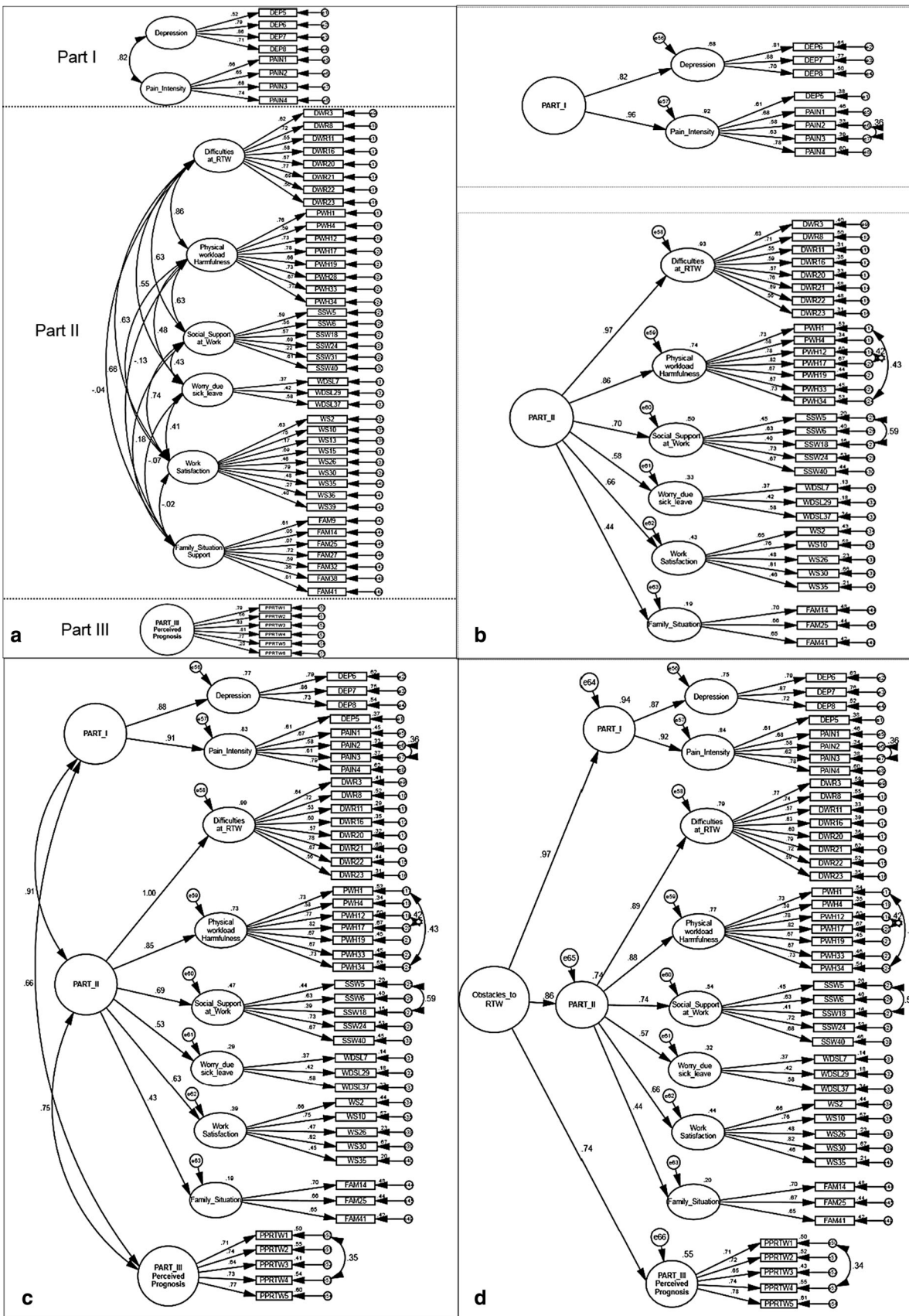
The orthogonal structure was initially proposed by Marhold and colleagues [5], who presented the

questionnaire in three parts. For the refinement of this model, 11 items in total were excluded (ten items in part II and one item in part III) and five correlations were included between the pairs of non-random errors of some items. These correlations were established considering the theoretical proximity/resemblance between items [38].

In Part I, the DEP5 item appeared problematic perhaps because of either theoretical formulation and/or redundancy, since it loaded meaningfully on both the Pain Intensity and the Depression subscales. In contrast to what has been suggested in a Swedish study [5], this item appears to be a stronger manifestation of Pain Intensity, than the Depression subscale, for the Brazilian sample.

It is worth emphasizing that the Family Situation and Support subscale in Part II presented as two different theoretical frameworks and, therefore, needed to be split—into a subscale on Family Situation, and another on Family Support. Furthermore, the items related to Family Support do not necessarily appear to be directly related to either the work environment or obstacles to RTW. Thus, they may have been misunderstood by the Brazilian sample that completed the questionnaire in our study. That is, those items may have led to response bias and, hence, the contribution of the Family Support subscale in relation to the central concept evaluated may have been compromised. It also presented a low explained variance (0.03%) for the Obstacles to RTW construct. Similarly, although the Family Situation subscale makes a significant contribution to the Obstacles to RTW concept, this significance is limited, as it most probably represents with the same weight for any life situation. Thus, since this subscale is not specific to the Brazilian sample under study, it may be considered as a type of 'white noise' [45].

One more consideration for Part II refers to the contribution of the Worry Due to Sick Leave subscale to the Obstacles to RTW construct which was low when compared with other subscales, such as Difficulties of Work Return or Physical Workload and Harmfulness. This limited contribution could be explained by the low number of items in this subscale and perhaps difficulties related to individual's interpretation on the meaning of "worry". This subscale has also presented with a number of other problematic issues, such as: low factor loadings, poor internal consistency and item cross-loadings. Methodologists suggest that for new scale development, at least three and preferably four measured variables for each common factor should be included [46]. In addition, others have recommended that factors with fewer than three items should not be retained [47], because they are generally weak and unstable [48]. Nonetheless, even with these issues, this subscale was kept in the model, because it presented a significant contribution to the assessed Obstacles to RTW construct and also, in order to keep the original theoretical framework formulated



**Fig. 1** Models of the Brazilian version of the ORTWQ: **a** orthogonal non-modified; **b** orthogonal second-order refined; **c** oblique refined and **d** oblique third order refined

by Marhold and colleagues [5]. Therefore, we believe that removing the Worry Due to Sick Leave subscale before assessing the adjustment of this model in other samples would be premature.

Results on the convergent validity of the factor structure by AVE presented values below the desirable cut-off for the majority of the subscales. Only Depression, Physical Workload and Harmfulness and Perceived Prognosis of Return to Work subscales demonstrated adequate levels of convergent validity. Discriminant validity limitations were also apparent for the subscales of Pain, Social Support at Work and Difficulties of Work Return. Limitations of the convergent and discriminant validities can be explained due to high correlations between items from the different subscales (i.e. cross-loadings) or even due to the high correlation ( $\rho$ ) between the subscales.

In the reliability evaluation, the internal consistency presented suitable values for all subscales except for the Worry Due to Sick Leave subscale. Evaluation of test–retest stability and of correlations between the ORTWQ subscales with the other questionnaires displayed adequate indices. Consideration should be given to the Family Situation subscale as there was not a tool available in order to evaluate this construct. Since we split the Family Situation and Support subscale, it was expected that the correlation of the Family Situation subscale would be low because the WHOQoL-Social Support subscale includes only items on support.

It is also worth mentioning, that the literature has discussed RTW or Obstacles/Facilitators to RTW as being a multidimensional theoretical model, where the subscales/parts

present some degree of relationship [8, 12]. Thus, we speculated the ORTWQ model would be better represented by an oblique model, which allows for correlations between the Parts I, II, and III.

Since the oblique model considers all relevant aspects from depression to perceived prognosis to return-to work, it would allow a more comprehensive and accurate assessment of the Obstacles to RTW construct. Our findings indeed confirmed this and we therefore recommend the oblique model proposal instead the orthogonal pattern. Furthermore, as we observed, all psychometric properties for the ORTWQ (structural, convergent and discriminant validities; concurrent validity, internal consistency and test–retest reliability) were similar regardless of which model, oblique and orthogonal, was used.

In order to calculate the ORTWQ score, previous Swedish study suggested [5] using a global score ranging from 0 to 330 points, by adding all of the scores from the nine subscales. However, we recommend calculating overall score using the regression weights matrix [38]. The algorithm is obtained for each population/sample individually and respects the cultural, demographics and social differences that may exist [17, 38]. Consequently, the algorithm generated in our study can also be used in any further Brazilian studies with similar characteristics to the inclusion criteria of our study, that is, 18 years old or older, presenting with a MSD and current work-absence related to their MSD or within the previous 60 days. Furthermore, the oblique modeling allowed us to estimate the score, not only for each subscale individually, but also for the higher hierarchical order factors. Thereby, this provides a targeted higher quality evaluation, to achieve of the originally intended goals of the questionnaire.

**Table 4** Correlational analysis between the ORTWQ subscales and other questionnaires: Non-modified and refined models of the Brazilian versions

ORTWQ scale		Questionnaire	r (Refined models)
Pain intensity	X	Pain scale (3 months)	0.53 (3 months)*
		(7 days)	0.56 (7 days)*
		WHOQoL (pain)	0.67*
Depression	X	WHOQoL (negative feelings)	0.58*
Difficulties at work return	X	SOPA (disability)	0.58*
		NFRS	0.62*
Physical workload and harmfulness	X	FABQ (work)	0.83*
Social support at work	X	JSS (support)	−0.71*
Work satisfaction	X	OSI (work satisfaction)	−0.58*
Family situation and support	X	WHOQoL (social support)	−0.22*(family situation)
Perceived prognosis of return to-work	X	FABQ (work—items 15,16)	0.67 (item 15)*
			0.66 (item 16)*
		SOPA (medical cure)	−0.51*

r = Spearman correlation coefficient

\*  $p$  value <0.01

**Table 5** Weights (W) of each item to be considered for calculating the overall score of ORTWQ’s subscales or of the Obstacles to Return-to-Work overall

Item	W <sup>a</sup>									TOHM <sup>b</sup>
	Subscales of ORTWQ (second order models)									
	PAIN	DEP	DWR	PHW	SSW	WDSL	WS	FAM	PPRTW	ORTW (overall)
PAIN1	0.115	0.036	0.029	0.007	0.006	0.006	0.004	0.004	0.005	0.084
PAIN2	0.06	0.019	0.015	0.004	0.003	0.003	0.002	0.002	0.003	0.043
PAIN3	0.069	0.022	0.017	0.004	0.004	0.004	0.002	0.003	0.003	0.051
PAIN4	0.173	0.054	0.044	0.01	0.009	0.01	0.006	0.007	0.008	0.115
DEP5	0.072	0.023	0.018	0.004	0.004	0.004	0.003	0.003	0.003	0.052
DEP6	0.025	0.19	0.024	0.006	0.005	0.005	0.003	0.004	0.004	0.061
DEP7	0.046	0.344	0.044	0.01	0.009	0.01	0.006	0.006	0.007	0.114
DEP8	0.019	0.141	0.018	0.004	0.004	0.004	0.002	0.003	0.003	0.043
DWR3	0.01	0.012	0.053	0.012	0.01	0.011	0.007	0.007	0.014	0.016
DWR8	0.017	0.02	0.087	0.019	0.017	0.018	0.011	0.012	0.023	0.019
DWR11	0.007	0.008	0.036	0.008	0.007	0.007	0.005	0.005	0.009	0.008
DWR16	0.008	0.009	0.04	0.009	0.008	0.008	0.005	0.005	0.01	0.009
DWR20	0.008	0.009	0.041	0.009	0.008	0.008	0.005	0.006	0.011	0.01
DWR21	0.017	0.02	0.089	0.019	0.017	0.018	0.011	0.012	0.023	0.019
DWR22	0.011	0.013	0.055	0.012	0.011	0.011	0.007	0.008	0.014	0.014
DWR23	0.011	0.014	0.059	0.013	0.012	0.012	0.008	0.008	0.015	0.014
PWH1	0.003	0.004	0.014	0.063	0.003	0.003	0.002	0.002	0.004	0.01
PWH4	0.003	0.003	0.012	0.055	0.003	0.003	0.002	0.002	0.003	0.009
PWH12	0.007	0.009	0.036	0.16	0.008	0.008	0.005	0.005	0.01	0.026
PWH17	0.009	0.011	0.043	0.192	0.009	0.009	0.006	0.006	0.012	0.031
PWH19	0.003	0.004	0.015	0.067	0.003	0.003	0.002	0.002	0.004	0.011
PWH33	0.002	0.003	0.012	0.052	0.002	0.003	0.002	0.002	0.003	0.009
PWH34	0.002	0.003	0.011	0.049	0.002	0.002	0.002	0.002	0.003	0.008
SSW5	0.001	0.001	0.005	0.001	0.04	0.001	0.001	0.001	0.001	0.004
SSW6	0.003	0.003	0.014	0.003	0.11	0.003	0.002	0.002	0.004	0.01
SSW18	0.001	0.001	0.003	0.001	0.022	0.001	0	0	0.001	0.002
SSW24	0.004	0.005	0.019	0.005	0.155	0.004	0.003	0.003	0.005	0.014
SSW40	0.003	0.003	0.014	0.003	0.112	0.003	0.002	0.002	0.004	0.011
WDSL7	0.001	0.001	0.006	0.001	0.001	0.071	0.001	0.001	0.002	0.004
WDSL29	0.002	0.002	0.007	0.002	0.002	0.086	0.001	0.001	0.002	0.005
WDSL37	0.003	0.003	0.013	0.003	0.003	0.15	0.002	0.002	0.004	0.009
WS2	0.002	0.003	0.012	0.003	0.002	0.003	0.126	0.002	0.003	0.008
WS10	0.003	0.003	0.014	0.003	0.003	0.003	0.147	0.002	0.004	0.01
WS26	0.001	0.001	0.005	0.001	0.001	0.001	0.052	0.001	0.001	0.003
WS30	0.004	0.005	0.019	0.004	0.004	0.004	0.198	0.003	0.005	0.013
WS35	0.001	0.001	0.004	0.001	0.001	0.001	0.047	0.001	0.001	0.003
FAM14	0.002	0.002	0.01	0.002	0.002	0.002	0.001	0.256	0.003	0.007
FAM25	0.002	0.002	0.01	0.002	0.002	0.002	0.001	0.251	0.003	0.007
FAM41	0.002	0.003	0.01	0.002	0.002	0.002	0.001	0.258	0.003	0.007
PPRTW1	0.002	0.002	0.012	0.003	0.003	0.003	0.002	0.002	0.1	0.022
PPRTW2	0.002	0.003	0.019	0.004	0.004	0.004	0.003	0.003	0.151	0.03
PPRTW3	0.002	0.002	0.014	0.003	0.003	0.003	0.002	0.002	0.109	0.025
PPRTW4	0.002	0.002	0.016	0.004	0.003	0.004	0.002	0.002	0.13	0.029

**Table 5** continued

Item	W <sup>a</sup>									
	Subscales of ORTWQ (second order models)									TOHM <sup>b</sup>
	PAIN	DEP	DWR	PHW	SSW	WDSL	WS	FAM	PPRTW	ORTW (overall)
PPRTW5	0.003	0.003	0.023	0.005	0.005	0.005	0.003	0.003	0.186	0.042

*PAIN* Pain, *DEP* depression, *DWR* difficulties to work return, *PHW* physical workload and harmfulness, *SSW* social support at work, *WDSL* worry due to sick-leave, *WS* work satisfaction, *FAM* family situation, *PPRTW* perceived prognosis of return-to-work

<sup>a</sup> Weights matrix of the model's regression

<sup>b</sup> Third-order hierarchical model

The analysis of the scores distribution of the ORTWQ and its subscales showed a substantial ceiling effect only for the Family Situation subscale. The Perceived Prognosis to Return-to-Work subscale also presented with a ceiling effect, but not as remarkable as the Family subscale. This may indicate that scores on the ORTWQ, and on its subscales, could be suitable for detecting changes on situations related to any obstacles to RTW. However, this parameter needs deeper longitudinal investigation.

### Strengths and Limitations

Any comprehensive discussion must acknowledge that our study has some limitations. The studied relied on self-reported data, which are subject to social desirability effects and other types of response bias (e.g., the selection bias as seen in any study involving volunteers). Also, the cross-sectional nature of this study did not accommodate an evaluation of the predictive validity of the ORTWQ. Additionally, the sample size, which although was adequate in our sample size calculation [48, 49], was at the limit of adequacy. Thus, we were not able to conduct the model's factorial invariance analysis in order to verify if the refined models would also be stable in independent samples.

Nonetheless these limitations are balanced by strengths in other areas. The use of confirmatory factor analysis allows a rigid evaluation of scales as compared to using the correlational evaluation between the ORTWQ and other questionnaires. The CFA results also provide stronger evidence of the scale's structural validity [50]. Other strengths of this study are the high response rate (94.3%), and the homogeneity of how the data were obtained, since the interviews were conducted by the same person.

### Relevance, Contributions and Implications

This study has resulted in key recommendations being made and was also able to evaluate a number of key issues that have been identified by other studies. For instance, researchers have recommended the conduct of studies that identify barriers and facilitators to RTW and to implement

appropriate measures [3]. Others have suggested that the ORTWQ was an instrument with the most promising results [7] but with a need to verify its psychometric properties in other populations and in larger samples [5, 7]. Recommendations from our study included reducing the number of the ORTWQ subscales, developing a range of interpretable cut-off scores, and assessing its ability to predict RTW as well as sick-leave duration [7].

The findings from this study can also be helpful to direct future research on the ORTWQ. As well, any of the specific items and some of its subscales may also contribute to the development and/or improvement of other tools. Moreover, the ORTWQ appears to have clinical potential, and might be used to direct more tailored RTW interventions based on the subscale's results. It may be used as a screening tool during rehabilitation to guide evaluation, goal setting and management of RTW for injured workers. Importantly, dissemination on the usability of the Brazilian version of the ORTWQ has been promoted in Brazilian conferences on Occupational Health. Also, the clinical settings, where this study was carried out, have been conducting some pilot projects in order to evaluate if ORTWQ may be useful for clinical indeed.

### Future Studies

In the work disability/RTW field it is essential that prospective studies be carried out in order to provide more reliable results, also in order to detect the causal relationship between the variables, detect the risk factors and to establish predictive values. Thus, the next step in the psychometric properties evaluation of the Brazilian version of the ORTWQ is to assess the predictive validity of this scale, using follow-up data on RTW.

We recommend that researchers evaluate the ORTWQ and its psychometric properties in different populations of workers with different characteristics, in order to continue improvement, refinement and testing of the ORTWQ.

This study has also identified that a questionnaire developed in one population does not always fit another population as originally intended. Some items may be

interpreted differently, or perhaps, may not be relevant at all. This leads to subscales requiring modification and the overall questionnaire scores being questioned, unless studies like this Brazilian study are conducted.

## Conclusions

The psychometric properties of the Brazilian version of the ORTWQ were evaluated and after refinement, the validity, reliability and floor/ceiling effects indexes were suitable when applied to a Brazilian sample of workers on sick-leave due to musculoskeletal disorders, although the factor structure presented some issues regarding convergent and discriminant validity. We suggest that the oblique pattern is the most comprehensive and accurate model to evaluate the construct of Obstacles to Return-to-Work and that the higher order hierarchical models can be used to produce an overall score in a more targeted way.

**Acknowledgments** We sincerely thank the Work Disability Prevention CIHR (Canadian Institutes of Health Research) Strategic Training Program, the Brazilian research agencies: Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) and the Conselho Nacional de Desenvolvimento Tecnológico (CNPq). A special thanks to Dr Susan E. Peters, for her diligent and meticulous contributions, advices and English editing of the final version of this study.

**Funding** DM was a trainee of Work Disability Prevention Program from 2012 to 2014 and has also funded by CAPES/CNPq to the Fellow of Doctoral Exchange Program by the Science without borders. Process number: 248514/2013-0.

## Compliance with Ethical Standards

**Conflict of interest** DM, NMCA, SHJ and JADBC declare that they have no conflict of interest.

**Ethical Approval** All procedures performed in this study were in accordance with ethical standards and with the 1964 Helsinki declaration and its later amendments. The research project of this study was submitted and approved by the University of Campinas Ethics Committee. (Protocol number: 143.694/2012 and amendments (607.710/2014)).

**Informed Consent** All participants signed two copies of consent form and the ORTWQ developers authorized the implementation of its cross-cultural adaptation and its use in Brazil.

## References

- Monteiro MS, Alexandre NM, Ilmarinen J, Rodrigues CM. Work ability and musculoskeletal disorders among workers from a public health institution. *Int J Occup Saf Ergon*. 2009;15(3):319–24.
- Punnett L, Wegman DH. Work-related musculoskeletal disorders: the epidemiologic evidence and the debate. *J Electromyogr Kinesiol*. 2004;14(1):13–23.
- Loisel P, Buchbinder R, Hazard R, Keller R, Scheel I, van Tulder M, et al. Prevention of work disability due to musculoskeletal disorders: the challenge of implementing evidence. *J Occup Rehabil*. 2005;15(4):507–24.
- Coutu MF, Durand MJ, Loisel P, Dupuis G, Gervais S. Measurement properties of a new quality of life measure for patients with work disability associated with musculoskeletal pain. *J Occup Rehabil*. 2005;15(3):295–312.
- Marhold C, Linton SJ, Melin L. Identification of obstacles for chronic pain patients to return to work: evaluation of a questionnaire. *J Occup Rehabil*. 2002;12(2):65–75.
- Iles RA, Davidson M, Taylor NF. Psychosocial predictors of failure to return to work in non-chronic non-specific low back pain: a systematic review. *Occup Environ Med*. 2008;65(8):507–17.
- Gray H, Adefolarin AT, Howe TE. A systematic review of instruments for the assessment of work-related psychosocial factors (Blue Flags) in individuals with non-specific low back pain. *Man Ther*. 2011;16(6):531–43.
- Pransky G, Gatchel R, Linton SJ, Loisel P. Improving return to work research. *J Occup Rehabil*. 2005;15(4):453–7.
- Loisel P, Durand MJ, Berthelette D, Vézina N, Baril R, Gagnon D, et al. Disability prevention—new paradigm for the management of occupational back pain. *Dis Manag Health Outcomes*. 2001;9(7):351–60.
- Linton SJ. Early identification and intervention in the prevention of musculoskeletal pain. *Am J Ind Med*. 2002;41(5):433–42.
- Gatchel RJ. Musculoskeletal disorders: primary and secondary interventions. *J Electromyogr Kinesiol*. 2004;14(1):161–70.
- Marois E, Durand MJ. Does participation in interdisciplinary work rehabilitation programme influence return to work obstacles and predictive factors? *Disabil Rehabil*. 2009;31(12):994–1007.
- Beaton D, Bombardier C, Guillemin F, Ferraz MB. Recommendations for the cross-cultural adaptation of the DASH & Quick DASH outcome measures. Institute for Work & Health, Toronto 2007, pp 1–45. <http://dash.iwh.on.ca/system/files/X-CulturalAdaptation-2007.pdf>. Accessed 01 Feb 2015.
- Streiner DL, Norman GR, Cairney J. Health measurements scales. A practical guide to their development and use. 5th ed. New York: Oxford University; 2015.
- Milani D, Hirayama MS, Souza AC, Alexandre NMC. Obstacles to return-to-work questionnaire: Cross-cultural adaptation to Brazilian context. *Cien Saude Colet*. 2016 (In press). [http://cienciaesaudecoletiva.com.br/artigos/artigo\\_int.php?id\\_artigo=15791](http://cienciaesaudecoletiva.com.br/artigos/artigo_int.php?id_artigo=15791).
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, & Joint Committee on Standards for Educational and Psychological Testing. Washington, DC: AERA; 2014.
- Terwee CB, Mokkink LB, Knol DL, Ostelo RW, Bouter LM, de Vet HC. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. *Qual Life Res*. 2011;21:651–7.
- Anastasi A. Psychological testing. 6th ed. New York: Macmillan Publishing Company; 1988.
- Whoqol Group. The development of the World Health Organization quality of life assessment instrument (the WHOQOL). In: Orley J, Kuyken W, editors. Quality of life assessment: international perspectives. Heidelberg: Springer; 1994. p. 41–57.
- Fleck MP, Leal OF, Louzada S, Xavier M, Chachamovich E, Vieira G, et al. Development of the Portuguese version of the OMS evaluation instrument of quality of life. *Rev Bras Psiquiatr*. 1999;21(1):19–28.
- Hartrick CT, Kovan JP, Shapiro S. The numeric rating scale for clinical pain measurement: a ratio measure? *Pain Pract*. 2003;3(4):310–6.



22. Sousa FF, Silva JA. The metric of pain: theoretical and methodological issues. *Rev Dor*. 2005;6(1):469–513.
23. Tait RC, Chibnall JT. Development of a brief version of the survey of pain attitudes. *Pain*. 1997;70(2–3):229–35.
24. Pimenta CAD, Kurita GP, da Silva EM, da Cruz DALM. Validity and reliability of the survey of pain attitudes (SOPA-28 items) in the Portuguese Language. *Rev Esc Enferm Usp*. 2009;43:1070–8.
25. van Veldhoven M, Broersen S. Measurement quality and validity of the “Need for Recovery Scale”. *Occup Environ Med*. 2003;60:3–9.
26. 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):131–9.
27. Waddell G, Newton M, Henderson I, Somerville D, Main CJ. A fear-avoidance beliefs questionnaire (FABQ) and the role of fear-avoidance beliefs in chronic low-back-pain and disability. *Pain*. 1993;52(2):157–68.
28. Abreu AM, Faria CD, Cardoso SM, Teixeira-Salmela LF. The Brazilian version of the fear avoidance beliefs questionnaire. *Cad Saude Publica*. 2008;24(3):615–23.
29. Karasek R, Brisson C, Kawakami N, Houtman I, Bongers P, Amick B. The job content questionnaire (JCQ): an instrument for internationally comparative assessments of psychosocial job characteristics. *J Occup Health Psychol*. 1998;3(4):322–55.
30. Alves MGD, Chor D, Faerstein E, Lopes CS, Werneck GL. Short version of the “job stress scale”: a Portuguese-language adaptation. *Rev Saude Publica*. 2004;38(2):164–71.
31. Cooper CL, Williams J. A validation-study of the OSI on a blue-collar sample. *Stress Med*. 1991;7(2):109–12.
32. Swan JA, Demoraes LFR, Cooper CL. Developing the occupational stress indicator (OSI) for Use in Brazil—a report on the reliability and validity of the translated OSI. *Stress Med*. 1993;9(4):247–53.
33. de Barros EN, Alexandre NM. Cross-cultural adaptation of the Nordic musculoskeletal questionnaire. *Int Nurs Rev*. 2003;50(2):101–8.
34. Hair JF, Black WC, Babin B, Anderson RE, Tatham RL. *Multivariate data analysis*. 7th ed. New York: Prentice Hall; 2009.
35. Kim KH. The relation among fit indices, power, and sample size in structural equation modeling. *Struct Equ Model*. 2005;12(3):368–90.
36. Guenther WC. Desk calculation of probabilities for distribution of sample correlation-coefficient. *Am Stat*. 1977;31(1):45–8.
37. Terwee CB, Bot SDM, de Boer MR, van der Windt DA, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol*. 2007;60(1):34–42.
38. Maroco J. *Analysis of Structural Equation: Theoretical Basis, Software & Applications*. (Title in Portuguese: *Análise de Equações Estruturais: Fundamentos teóricos, Software e Aplicações*). 2nd ed. Pêro Pinheiro: Report Number; 2014.
39. Kline RB. *Principles and practice of structural equation modeling*. 3rd ed. New York: Guilford; 2011.
40. Schultz IZ, Stowell AW, Feuerstein M, Gatchel RJ. Models of return to work for musculoskeletal disorders. *J Occup Rehabil*. 2007;17(2):327–52.
41. Wasiaik R, Young AE, Roessler RT, McPherson KM, van Poppell MN, Anema JR. Measuring return to work. *J Occup Rehabil*. 2007;17(4):766–81.
42. Williams B, Onsmann A, Brown T. *Exploratory factor analysis: a five-step guide for novices*. JEPHC. 2010;8(3):1–13.
43. Fornell C, Larcker DF. Evaluating structural equation models with unobservable variables and measurement error. *J Mark Res*. 1981;18(1):39–50.
44. Cohen J. *Statistical power analysis for the behavioral sciences*. Hillsdale: Lawrence Erlbaum Associates Inc; 1988.
45. Kuo H-H. *Probability and stochastics series: white noise distribution theory—Book 5*. Boca Raton: CRC Press; 1996.
46. Fabrigar LR, Wegener DT, MacCallum RC, Strahan EJ. Evaluating the use of exploratory factor analysis in psychological research. *Psychol Methods*. 1999;4(3):272–99.
47. Tabachnick BG, Fidell LS. *Using multivariate statistics*. 6th ed. Boston: Allyn & Bacon; 2007.
48. Costello AB, Osborne JW. Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis. *Pract Assess Res Eval*. 2005;10:1–9.
49. Campbell P, Bishop A, Dunn KM, Main CJ, Thomas E, Foster NE. Conceptual overlap of psychological constructs in low back pain. *Pain*. 2013;154(9):1783–91.
50. Martus P, Jakob O, Rose U, Seibt R, Freude G. A comparative analysis of the work ability index. *Occup Med*. 2010;60(7):517–24.