Psychometric Properties of Self-Reported Instruments for Occupational Balance: A COSMIN-Based Systematic Review

Sangmi Park [corresponding author] sangmipark@yonsei.ac.kr

https://orcid.org/0000-0002-3029-1772

Department of Occupational Therapy, College of Software and Digital Healthcare Convergence, Yonsei University, Republic of Korea; 443 Backun Hall, 1 Yonseidae-gil, Wonju-si, Gangwon-do, Republic of Korea, Zip code: 26493

Chang Dae Lee

https://orcid.org/0000-0001-7606-4092

Human Engineering Research Laboratories, Department of Veterans Affairs Pittsburgh Healthcare System, Pittsburgh, PA, USA; Human Engineering Research Laboratories, Department of Rehabilitation Science and Technology, University of Pittsburgh, Pittsburgh, PA, USA; 6425 Penn Avenue, Suite 400, Pittsburgh, PA, USA, Zip code: 15206

See final published article at Sangmi P & Chang Dae L. (2023) Psychometric properties of self-reported instruments for occupational balance: A COSMIN-based systematic review, *Journal of Occupational Science*, DOI: 10.1080/14427591.2023.2243281

Psychometric properties of self-reported instruments for occupational balance: A COSMIN-based Systematic Review

ABSTRACT

Introduction: Instruments being used currently to evaluate Occupational Balance (OB) are all self-report outcome measures which demand robust psychometric properties for precise assessment. This study aimed to systematically review the psychometric properties of the current OB instruments.

Methods: For this systematic review, COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) guidelines and Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 were adopted. Peer-reviewed articles published between January 2010 and February 2021 were searched for in PubMed, Cochrane, Web of Science, and PsycINFO databases. Search terms were full names and abbreviations of the OB instruments: Life Balance Inventory (LBI), Meaningful Activity Wants and Needs Assessment (MAWNA), Occupational Balance Questionnaire (OBQ), and Occupational Balance-Questionnaire (OB-Quest).

Results: Of the 400 articles identified in a systematic database search, 12 studies from ten articles were analysed. Five instruments, including the revised version of the OBQ (i.e. OBQ11), were analysed within the COSMIN taxonomy. The LBI demonstrated ‘high’ evidence for insufficient structural validity. The others demonstrated ‘moderate’ or ‘high’ evidence for indeterminate structural validity. Less than sufficient structural validity led to indeterminate internal consistency across five instruments, and their quality of evidence varied from ‘very low’ to ‘high’. The MAWNA and OBQ demonstrated ‘very low’ evidence for sufficient reliability.

See final published article at Sangmi P & Chang Dae L. (2023) Psychometric properties of self-reported instruments for occupational balance: A COSMIN-based systematic review, *Journal of Occupational Science*, DOI: 10.1080/14427591.2023.2243281

Conclusion: OB instruments with psychometric properties of sufficient overall ratings supported by 'high' quality of evidence are unavailable. Further studies are required to clarify the concept of OB, to examine more diverse psychometrics of OB instruments, and to consolidate their robustness.

Keywords: psychometrics, life balance, occupational therapy, patient-reported outcome measure

Introduction

Occupational balance (OB), commonly known as 'life balance', 'lifestyle balance', 'work-life balance', or 'work-leisure balance' (Matuska, 2012; Wagman, Håkansson, Jacobsson, et al., 2012; Wilcock & Hocking, 2015), refers to how individuals organize their daily life harmoniously across various occupations (Wagman, Håkansson, & Björklund, 2012; Matuska, 2012). These occupations are variously categorized as work, play, leisure, rest and sleep, activities of daily living, instrumental activities of daily living, education, and social participation (Backman, 2004; Eklund et al., 2017; Jonsson & Persson, 2006; Stamm et al., 2009; Wagman, Håkansson, & Björklund, 2012; Wilcock et al., 1997). OB is a significant topic of interest in occupational science and occupational therapy because these disciplines, respectively, focus on understanding human occupation and improving people's participation in meaningful and purposeful occupations (Sainburg et al., 2017; Wilcock, 1991).

Many researchers have been interested in measuring OB to explore its association with health-related outcomes, including boredom, fatigue, exhaustion, burnout, stress, well-being, and

See final published article at Sangmi P & Chang Dae L. (2023) Psychometric properties of self-reported instruments for occupational balance: A COSMIN-based systematic review, *Journal of Occupational Science*, DOI: 10.1080/14427591.2023.2243281

life satisfaction (Clouston, 2014, 2015; Jonsson & Persson, 2006; Lee et al., 2021; Park et al., 2021). As there is not a gold-standard instrument for OB nor consensus about how to define it, the construct of OB has been measured using diverse methods, depending on the researchers (Dür et al., 2015). Earlier research has focused mainly on time-use studies using instruments, such as the Occupational Questionnaire (OQ; Kielhofner, 2007), NIH Activity Record (NIH ACTRE; Gerber & Furst, 1992), and structured observation and report techniques (SORT; Rintala et al., 1984).

Time-use studies focus primarily on understanding occupational patterns (i.e., issues with pacing and prioritising occupations). Understanding individual's patterns of occupation based on the results of time-use assessment is useful to identify the negative and positive factors for health and well-being among daily routines (Gerber & Furst, 1992). However, the results provide limited information to understand the balance among occupations. Results of OB assessment reflect subjective satisfaction with balance among diverse daily occupational engagement which is influenced by the roles, values, or interests of the person (Håkansson & Matuska, 2010; Wagman, Håkansson, Jacobsson, et al., 2012). Although appropriateness for time-use is one contributing factor affecting the satisfaction (Eklund et al., 2010), satisfaction with time-use per se is not perceived OB.

Backman and Forhan (2017) differentiated OB instruments from time-use instruments and listed six OB instruments: Life Balance Inventory (LBI; Matuska, 2012), Meaningful Activity Wants and Needs Assessment (MAWNA; Eakman, 2015), Occupational Balance Questionnaire (OBQ; Wagman & Håkansson, 2014a), Occupational Balance-Questionnaire (OB-Quest; Dür et al., 2014), Inter-goal Relations Questionnaire (IRQ; Riediger et al., 2005), and

See final published article at Sangmi P & Chang Dae L. (2023) Psychometric properties of self-reported instruments for occupational balance: A COSMIN-based systematic review, *Journal of Occupational Science*, DOI: 10.1080/14427591.2023.2243281

Personal Projects Analysis (PPA; Little, 1983). These instruments are distinctive from time-use instruments in that the results of these instruments indicate the levels of perceived OB directly. Among these, however, the IRQ and PPA cover limited aspects of occupation (Dür et al., 2015). The IRQ examines personal goals and the intensity of goal pursuit. Similarly, PPA focuses on personal projects and ranks the importance of each project. Specifically, these two instruments are focused on the personal goals or personal projects which are the focus of the respondent's life-span-specific occupations. Furthermore, the results of the IRQ and PPA omit subjective satisfaction with balance among occupational engagement. Hence, it is challenging to view the IRQ and PPA as measures of OB. In 2012, OB instruments began to be introduced which were developed by occupational science and occupational therapy researchers to measure the extent of perceived balance among diverse occupational engagements. Currently, LBI, MAWNA, OBQ, and OB-Quest have been suggested as OB measures (Håkansson et al., 2020). These four instruments were developed as self-reported instruments for measuring perceived OB (Backman & Forhan, 2017). Items in these instruments ask satisfaction with participation in each of daily occupations (Dür et al., 2014; Eakman, 2015; Matuska, 2012) or ask the degree of perceived balance across various daily occupations to measure perceived OB (Dür et al., 2014; Wagman & Håkansson, 2014a). Levels of perceived OB can be determined by the person who participates in occupations himself/herself. Given that perceived OB can be measured most accurately by oneself, self-reported outcome measures are the most preferred type of instrument for OB assessment. Robust psychometric properties are essential for self-reported outcome measures. This is because the results of self-reported assessments are predominantly influenced by respondents' feelings and experiences (Fiore Jr et al., 2018; Kroenke et al., 2015).

The psychometric properties of each of these four OB instruments have been

See final published article at Sangmi P & Chang Dae L. (2023) Psychometric properties of self-reported instruments for occupational balance: A COSMIN-based systematic review, *Journal of Occupational Science*, DOI: 10.1080/14427591.2023.2243281

independently examined and reported in previous studies. However, whether the reported psychometric properties are sufficient as much as they meet the criteria suggested for self-reported outcome measures has not been explored (Prinsen et al., 2018; Reeve et al., 2013). Also, the quality of evidence for the reported psychometric properties (e.g. whether there was any bias in the methodology or whether similar results were found from repeated studies) had not been analysed. Therefore, this study aimed to examine the psychometric properties of OB instruments in terms of sufficiency and the quality of evidence by synthesising the previous reports to determine whether the current OB instruments have robust psychometric properties as self-report instruments.

Methods

This systematic review adopted the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) methodology for systematic reviews of Patient Reported Outcome Measures (PROMs). These guidelines are pertinent for the analysis of the psychometric properties of a self-reported instrument (Prinsen et al., 2018). The methods for evaluating the overall ratings and quality of evidence for the psychometric properties were adopted for this study. The guidelines can be applied to systematically review the psychometric properties of one specific instrument (Ohno et al., 2021; Pike et al., 2018). Furthermore, they can be applied to systematically review the psychometric properties across more than two instruments which measure one specific construct (Harrison et al., 2022; Jewell et al., 2019; Shoman et al., 2021; Soh et al., 2021).

We followed the Preferred Reporting Items for Systematic reviews and Meta-Analyses

See final published article at Sangmi P & Chang Dae L. (2023) Psychometric properties of self-reported instruments for occupational balance: A COSMIN-based systematic review, *Journal of Occupational Science*, DOI: 10.1080/14427591.2023.2243281

2020 (PRISMA, 2020; Page et al., 2021). All authors participated in designing this review study. The protocol for this review was registered with PROSPERO (registration number: CRD42021211503; https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42021211503).

The entire procedure, from the study selection process to data synthesis, was independently conducted by two researchers. The results of each step were compared and combined. Where there were conflicts in the results, an agreement was reached through discussion between the two researchers. Discussion was necessary for the risk of bias analysis step. In this step, the inter-rater reliability was 75.0%. The disagreement of the results was due to the assumable but not specified methods and results in the included studies. After discussion, analyses were limited to the specified methods and results.

Information Sources and Search Strategy

The data search was conducted from 23 to 25 February, 2021, by a medical librarian with seven years of experience in data search for systematic reviews. Search terms were selected by the authors and the databases to be searched were discussed with the librarian. Data search using search terms was independently conducted by a librarian. PubMed, Cochrane, Web of Science, and PsycINFO were searched using keywords containing the full names and abbreviations of the four OB instruments (Appendix 1). A sensitive filter designed by the COSMIN research group to search for studies with psychometric properties (Terwee et al., 2009) was applied.

The year of publication was set based on the date of the first developed instrument (i.e., LBI, submitted in 2011 and published in 2012; Matuska, 2012) among the four OB instruments.

See final published article at Sangmi P & Chang Dae L. (2023) Psychometric properties of self-reported instruments for occupational balance: A COSMIN-based systematic review, *Journal of Occupational Science*, DOI: 10.1080/14427591.2023.2243281

To ensure the inclusion of studies published in 2012 and studies published before 2012 due to the time difference of publication, studies published between 1 January 2010 and 25 February 2021 were searched.

Eligibility Criteria

Among the six OB instruments suggested by Backman and Forhan (2017), the authors were interested in four: LBI, MAWNA, OBQ, and OB-Quest. This selection is supported by a previous study published in 2020 by Håkansson et al. (2020). In addition, the revised version of the OBQ (i.e., OBQ11; Håkansson et al., 2020) was added. The OBQ11 was recognised during the screening process of this study, and the authors consented to accommodate the OBQ11 as an instrument to measure perceived OB. Inclusion criteria were updated to treat the OBQ and OBQ11 as independent instruments. This was because differences in the psychometrics of the OBQ and OBQ11 were expected due to the differences in the number of items and the scale of the two instruments. After the inclusion criteria were updated, data were rescreened accordingly.

The inclusion criteria were updated as follows: 1) development studies of the five instruments: LBI, MAWNA, OB-Quest, two versions of OBQ (i.e. original OBQ; OBQ11), or studies that applied any one of these instruments; 2) studies that measured the OB of adults aged 18 years or older; 3) studies that reported any of the five psychometric properties (i.e., structural validity, internal consistency, cross-cultural validity/measurement invariance, reliability, and measurement error) defined by the COSMIN taxonomy (Mokkink et al., 2010); and 4) studies published in peer-reviewed journals in English since 2010. Studies we could not access in full-text were excluded.

See final published article at Sangmi P & Chang Dae L. (2023) Psychometric properties of self-reported instruments for occupational balance: A COSMIN-based systematic review, *Journal of Occupational Science*, DOI: 10.1080/14427591.2023.2243281

Study Selection Process

After eliminating duplicates, the two researchers independently screened the titles, abstracts, and full texts to check the eligibility of the searched articles. We used the review production platform Covidence (https://www.covidence.org/) for efficient and precise screening and selection.

Data Collection Process

We used an extraction format modified based on the templates included in the COSMIN guidelines (Prinsen et al., 2018). Information on the study sample and psychometric properties (e.g., magnitudes of related coefficients) was extracted from the included studies.

Risk of Bias Assessment

The COSMIN bias checklist (Mokkink et al., 2018) was used to rate the methodological quality. The checklist questions asked whether psychometric properties were examined under proper testing conditions regarding study design (e.g., Was an internal consistency statistic calculated for each unidimensional scale or subscale separately?), sample size (e.g., Was the sample size adequate? For factor analysis, seven times the number of items and more than 100, For Rasch 1PL model, more than 200), and statistical methods (e.g. Were proper statistics calculated? For continuous scores, Cronbach's alpha or omega, For dichotomous scores, Cronbach's alpha or KR-20), without any significant flaws in the study methodology.

The checklist was modularly applied to the corresponding examination according to each

See final published article at Sangmi P & Chang Dae L. (2023) Psychometric properties of self-reported instruments for occupational balance: A COSMIN-based systematic review, *Journal of Occupational Science*, DOI: 10.1080/14427591.2023.2243281

psychometric property. The results were obtained using a two-step procedure. First, each question in the applied module was rated as very good, adequate, doubtful, inadequate, or inapplicable. Second, the lowest score within the module was adopted as the result of the risk of bias in related psychometric properties.

Synthesis Methods: Overall Rating and Quality of Evidence

Extracted data on psychometric properties from the included studies was evaluated independently and then those results were synthesized into an overall rating and quality of evidence for each psychometric property per instrument. The overall rating is the representative value of the ratings. Quality of evidence indicates the level of credibility for the overall rating.

Ratings were evaluated according to the criteria for good psychometric properties in the COSMIN guidelines. The magnitude of the coefficients per psychometric property was evaluated as sufficient (+), insufficient (−), or indeterminate (?). The ratings were synthesised into an overall rating per psychometric property. The overall ratings were sufficient (+), insufficient (−), indeterminate (?), or inconsistent (±). The overall ratings were determined by the 75% rule for the ratings. For example, an overall rating in structural validity of instrument A was rated 'sufficient (+)' when at least 75% of the ratings of structural validity in instrument A were sufficient, 'insufficient (−)' when at least 75% of the ratings were insufficient, and 'indeterminate (?)' when all ratings were indeterminate. The overall rating was rated as 'inconsistent (±)' when the ratings were mixed including sufficient and insufficient, and the reasons for the mixed results were not fully explained.

The quality of evidence for the overall rating was graded based on the modified Grading

See final published article at Sangmi P & Chang Dae L. (2023) Psychometric properties of self-reported instruments for occupational balance: A COSMIN-based systematic review, *Journal of Occupational Science*, DOI: 10.1080/14427591.2023.2243281

of Recommendations Assessment, Development, and Evaluation (GRADE) approach (Prinsen et al., 2018). The quality of evidence was graded as 'high', 'moderate', 'low', or 'very low', taking into account four factors: the risk of bias, inconsistency (i.e., unexplained inconsistency in the overall rating), imprecision (i.e., small size of the total sample), and indirectness (i.e., irrelevant population). For example, 'high' quality of evidence for 'sufficient (+)' reliability of instrument A means that the degree of confidence on identicalness between the result (i.e., sufficient reliability) and the true measurement of A is high.

Results

Study Selection

We collected 400 references from four databases (PubMed, 82; Cochrane, 34; Web of Science, 204; PsycINFO, 80). After removing 66 duplicates, 334 articles were screened for titles and abstracts (Figure 1). After excluding 304 articles, the full text of 30 articles was assessed for eligibility. Ten articles were included in the final analysis. Two out of the 10 articles contained two sub-studies. According to the COSMIN guidelines, the sub-study needed to be separately analysed if it met the inclusion criteria. Hence, 12 studies were analysed.

See final published article at Sangmi P & Chang Dae L. (2023) Psychometric properties of self-reported instruments for occupational balance: A COSMIN-based systematic review, *Journal of Occupational Science*, DOI: 10.1080/14427591.2023.2243281

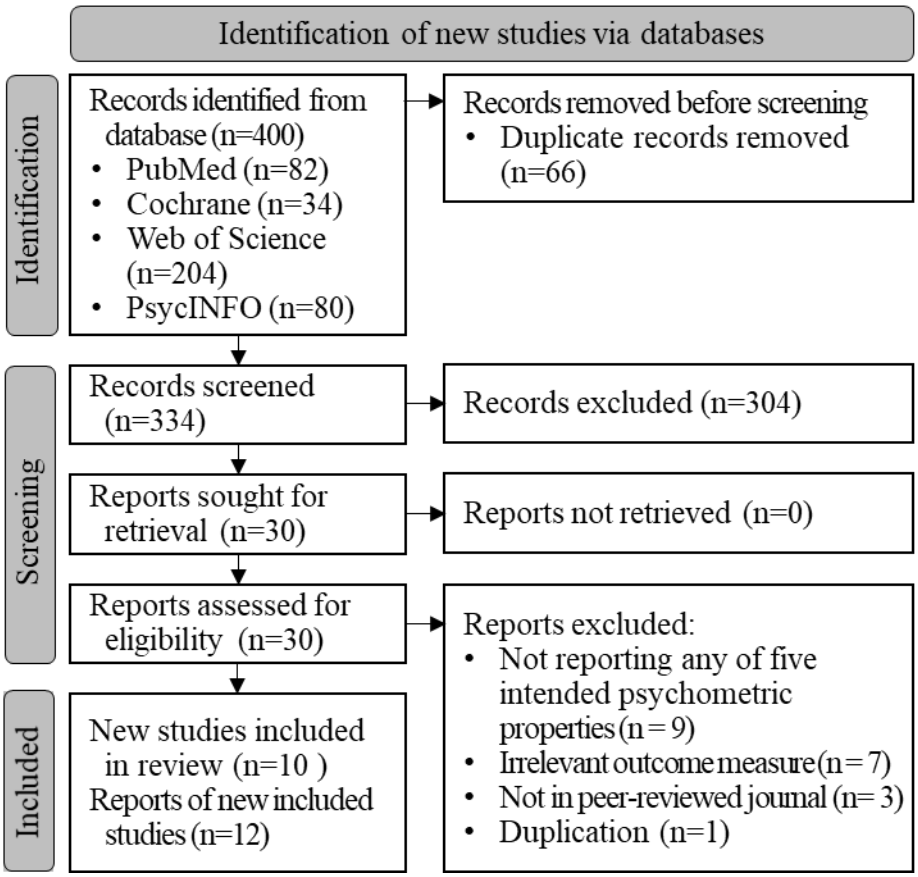

Fig. 1. Flow diagram for selecting studies according to PRISMA 2020

Note. Twelve studies (reports) from ten articles (studies) were included.

Study Characteristics

As the OBQ and OBQ11 were evaluated separately, the results of this study dealt with five OB instruments. The study characteristics, classified by the five instruments, are listed in Table 1. General information on the five OB instruments, domains, and scoring systems is summarised in Appendix 2.

See final published article at Sangmi P & Chang Dae L. (2023) Psychometric properties of self-reported instruments for occupational balance: A COSMIN-based systematic review, *Journal of Occupational Science*, DOI: 10.1080/14427591.2023.2243281

Results of Individual Studies: Rating of psychometric properties and Risk of Bias

Extracted data on study sample, the coefficients of psychometric properties, ratings according to the criteria for good psychometric properties, and methodological quality (i.e., risk of bias) across the included studies were presented in Table 1. As a result of analysing 12 studies in accordance with COSMIN guidelines, structural validity, internal consistency, and reliability (i.e., test-retest reliability) were analysed among the five types of psychometric properties which were considered for analysis in this study. The psychometric properties commonly examined across the five instruments were structural validity and internal consistency. Reliability was examined for the MAWNA and OBQ.

Table 1. Study sample, methodological quality, and psychometric properties across the included studies

Author (Year)	Study sample		Structural validity		Internal consistency		Reliability	
	Mean Age (SD) / Range	Gender (%F)	*n*	Risk of bias / Coefficient (Rating) [a]	*n*	Risk of bias / Coefficient (Rating) [b]	*n*	Risk of bias / Coefficient (Rating) [c]
LBI								
Matuska (2012)	42.8 (- [d]) / 19-90	90	282	Doubtful / - [e] (?)	282	Inadequate / α =0.89 (?)		
Matuska (2012)	41.1 (- [d]) / 18-90	55	458	Very good / RMSEA =0.8, CFI=0.7 (-)	458	Inadequate / α =0.97 (?)		
Farber et al (2020)	50.1 (11.5) / 25-71	94			108	Very good / α=0.89 (overall), α=0.57-0.73 (domain) (?)		
MAWNA								
Eakman (2015)	31.1 (9.3) / - [d]	64.8	250	Adequate / - [e] (?)	250	Inadequate / α=0.95 (?)	49	Doubtful / ICC=.89 (+)
OBQ								
Wagman & Håkansson (2014a)	49.6 (11.3) / - [d]	70.3			67	Very good / α=0.94 (?)	67	Doubtful / Spearman Rho correlation=0.93 (+)
Håkansson et al (2020)	53.0 (12.0) / - [d]	68.0	168	Adequate / - [e] (?)	168	Very good / PSI=0.93 (?)		
Wagman & Håkansson (2014b)	M: 39.8 (11.4) / - [d] F: 42.4 (12.8) / - [d]	63.4			153	Very good / α=0.92 (?)		
Ortiz-Rubio et al (2020)	49.5 (8.4) / - [d]	100			100	Very good / α=0.89 (?)		
Wagman et al (2020)	62.0 (12.0) / 24-81	73.0			682	Very good / α=0.95 (?)		
OBQ11								
Lexén et al (2020)	45 .0 (11.0) / - [d]	94			3658	Very good / α=0.94 (?)		
Håkansson et al (2020)	50.0 (13.0) / 20-89	64	207	Adequate / - [e] (?)	207	Very good / PSI=0.92 (?)		
OB-Quest								
Dür et al (2014)	- [f] (- [d]) / 27-65	81.7	251	Very good / - [e] (?)	251	Doubtful / α=0.57 (?)		

Note. Included twelve studies were sorted by the name of instrument. Not reported or not applicable psychometric properties (i.e., cross-cultural validity/measurement invariance and measurement error) were not presented in the table.
(+), (-), (?) means "sufficient", "insufficient", "indeterminate", respectively.
α=Cronbach's alpha; CFI=Comparative Fit Index; F=Female; ICC=Intraclass Correlation Coefficient; LBI=Life Balance Inventory; M= Male; MAWNA=Meaningful Activity Wants and Needs Assessment; OB-Quest=Occupational Balance Questionnaire; OBQ=Occupational Balance Questionnaire; PSI=Person Separation Index; RMSEA=Root Mean Square of Error of Approximation; SD=Standard Deviation.
[a] Criteria for good structural validity: (+)=CTT: CFA: CFI or TLI or comparable measure >0.95 or RMSEA<0.06 or SRMR<0.08, IRT/Rasch: No violation of unidimensionality AND No violation of local independence AND No violation of monotonicity AND adequate model fit, (-)=Criteria for '+' not met , (?)=CTT: Not all information for '+' reported, IRT/Rasch: Model fit not reported.
[b] Criteria for good internal consistency: (+)=At least low evidence for sufficient structural validity AND Cronbach's alpha(s) ≥0.70 for each unidimensional scale or subscale, (-)=At least low evidence for sufficient structural validity AND Cronbach's alpha(s) < 0.70 for each unidimensional scale or subscale, (?)=Criteria for "At least low evidence for sufficient structural validity" not met.
[c] Criteria for good reliability: (+)=ICC or weighted Kappa≥0.70, (-)=ICC or weighted Kappa < 0.70, (?)=ICC or weighted Kappa not reported.
[d] Information was not provided.
[e] Reported estimates were not ratable from within COSMIN guidelines.
[f] Mean age according to subgroups was reported separately (Healthy group: 38, Rheumatoid arthritis group: 59, Systemic lupus erythematosus group: 46).

See final published article at Sangmi P & Chang Dae L. (2023) Psychometric properties of self-reported instruments for occupational balance: A COSMIN-based systematic review, *Journal of Occupational Science*, DOI: 10.1080/14427591.2023.2243281

Structural validity across the five instruments was reported in six studies (Table 1). Except for the insufficient rating of the structural validity of the LBI in one study, the ratings of the structural validity of the five instruments were indeterminate. This was attributed to the absence or partial reporting of information on the model fit of the construct in each study. Regarding the risk of bias, one study using LBI was rated as doubtful for the question, 'For Item Response Theory (IRT) or Rasch: Does the chosen model fit the research question?' owing to the insufficient description of the model, although the research question was clearly described. The other five studies demonstrated adequate or very good levels of methodological quality.

Internal consistency across the five instruments was reported in all the included studies (Table 1). The magnitude of the coefficient of internal consistency reported in 11 studies, except one study (Farber et al., 2020), met the criteria of the guidelines (i.e., Cronbach's alpha ≥ 0.70). However, the ratings of internal consistency for all five instruments were indeterminate. This was because, to be evaluated as sufficient with internal consistency according to COSMIN guidelines (Prinsen et al., 2018), the magnitude of the coefficient of internal consistency has to meet the criteria in addition to sufficient structural validity. Regarding the risk of bias, eight out of the 12 studies showed a very good methodological quality. These were all studies using the OBQ and OBQ11 and one using the LBI. The risk of bias in three studies (two studies using LBI and one using MAWNA) was rated as inadequate. The risk of bias in one study using OB-Quest was rated as doubtful for the question, 'Was an internal consistency statistic calculated for each unidimensional (sub)scale separately?'.

Reliability was reported in two studies, one each for MAWNA and OBQ (Table 1). The reliability ratings for both the MAWNA and OBQ were sufficient. Regarding the risk of bias,

See final published article at Sangmi P & Chang Dae L. (2023) Psychometric properties of self-reported instruments for occupational balance: A COSMIN-based systematic review, *Journal of Occupational Science*, DOI: 10.1080/14427591.2023.2243281

both studies showed lower than adequate methodological quality. The risk of bias in both studies was rated as doubtful for the question, 'Were patients stable in the interim period on the construct to be measured?'.

Results of Syntheses: Overall Rating and Quality of Evidence

The summary of the coefficients, overall rating, and quality of evidence per psychometric property and instrument are presented in Table 2.

LBI demonstrated a 'high' quality of evidence for 'insufficient (−)' structural validity. This means, synthesized results of the magnitudes of the coefficients of the structural validity of LBI from included studies was lower than the standard, and this insufficient overall rating was supported by high quality evidence. LBI demonstrated a 'low' quality of evidence for 'indeterminate (?)' internal consistency. MAWNA demonstrated a 'moderate' quality of evidence for 'indeterminate (?)' structural validity, a 'very low' quality of evidence for 'indeterminate (?)' internal consistency, and a 'very low' quality of evidence for 'sufficient (+)' reliability. OBQ demonstrated a 'moderate' quality of evidence for 'indeterminate (?)' structural validity, a 'high' quality of evidence for 'indeterminate (?)' internal consistency, and a 'very low' quality of evidence for 'sufficient (+)' reliability. OBQ11 demonstrated a 'moderate' quality of evidence for 'indeterminate (?)' structural validity, and a 'high' quality of evidence for 'indeterminate (?)' internal consistency. OB-Quest demonstrated a 'high' quality of evidence for 'indeterminate (?)' structural validity, and a 'low' quality of evidence for 'indeterminate (?)' internal consistency.

See final published article at Sangmi P & Chang Dae L. (2023) Psychometric properties of self-reported instruments for occupational balance: A COSMIN-based systematic review, *Journal of Occupational Science*, DOI: 10.1080/14427591.2023.2243281

Table 2. Overall rating and quality of evidence for psychometric properties of each instrument

Instrument	Structural validity			Internal consistency			Reliability		
	Summary or pooled result	Overall rating	Quality of evidence	Summary or pooled result	Overall rating	Quality of evidence	Summary or pooled result	Overall rating	Quality of evidence
LBI	4 factors	(-) [a]	High	α=0.89-0.97	(?)	Low	n/a	n/a	n/a
MAWNA	n/a	(?)	Moderate	α=0.95	(?)	Very Low	ICC=.89	(+)	Very Low
OBQ	n/a	(?)	Moderate	α=0.89-0.95	(?)	High	ρ =0.93	(+)	Very Low
OBQ11	n/a	(?)	Moderate	α=0.92-0.94	(?)	High	n/a	n/a	n/a
OB-Quest	n/a	(?)	High	α=0.57	(?)	Low	n/a	n/a	n/a

Note. Not reported or not applicable psychometric properties (i.e., cross-cultural validity/measurement invariance and measurement error) were not presented in the table.

(+), (-), (?) means "sufficient", "insufficient", "indeterminate" respectively. α=Cronbach's alpha; ICC=Intraclass Correlation Coefficient; LBI=Life Balance Inventory; MAWNA=Meaningful Activity Wants and Needs Assessment; OB-Quest=Occupational Balance Questionnaire; OBQ=Occupational Balance Questionnaire; ρ= Spearman Rho correlation

[a] Following the COSMIN guideline, (?) and (-) in structural validity of LBI were summarized into (-) which is the less biased study.

Discussion

In this study, psychometric properties of the five OB instruments were analysed according to the COSMIN guidelines. None of the five instruments exhibited 'high' evidence for 'sufficient (+)' psychometric properties in accordance with the standards of COSMIN guidelines. Given that the instruments have been developed as self-report instruments, robustness of psychometric properties of the current OB instruments needs to be strengthened. The results of this study provide useful information on psychometrics of the five instruments which are currently applicable to measure perceived OB in research and practice.

In this study, inconsistencies between the estimates reported in studies and standards of COSMIN guidelines led to the indeterminate ratings from within the criteria for good psychometric properties, and finally led to the 'indeterminate (?)' overall ratings of the structural validity of MAWNA, OBQ, OBQ11, and OB-Quest. This indicates a lack of evidence to evaluate their structural validity as sufficient or insufficient within COSMIN guidelines. Nonetheless, these four instruments can have some degree of structural validity. This interpretation is supported by an unbiased study (i.e., adequate or very good level of methodological quality) in

See final published article at Sangmi P & Chang Dae L. (2023) Psychometric properties of self-reported instruments for occupational balance: A COSMIN-based systematic review, *Journal of Occupational Science*, DOI: 10.1080/14427591.2023.2243281

which structural validity was examined (Prinsen et al., 2018).

LBI demonstrated a ‘high’ quality of evidence for ‘insufficient (-)’ structural validity in this study. This insufficient structural validity could be due to the characteristics of the concept of OB, which respects individuals’ diverse daily patterns and dissimilar values for the same activities. The LBI consists of 53 items covering most of modern Western life’s daily activities (Matuska, 2012). The respondents could skip rating items they had not participated in within the last 4 weeks or if the activities identified were not meaningful to them. As the excluded items can be disparate by respondents, a four-factor structure based on the life balance model could not fit the collected data.

According to the COSMIN guidelines, ‘high’ quality of evidence can be graded if there are no deductible points in any of four factors: risk of bias, inconsistency, imprecision, and indirectness (Prinsen et al., 2018). The reason for downgrading as ‘low’ or ‘very low’ quality of evidence for internal consistency and reliability of the instruments explored in this study was mainly because of imprecision and risk of bias. More studies reporting the psychometric properties of OB instruments is needed to make up for this imprecision. Regarding the risk of bias, not having internal consistency statistics calculated for each subscale, a test-retest interval other than two weeks, or insufficient description of the related methodology were found to be the factors that lower methodological quality. In previous review studies within occupational therapy, these issues were detected as reasons for deducted points in the risk of bias evaluation, eventually leading to a downgrading in the quality of evidence (Pike et al., 2018; Yuen & Austin, 2014). Overcoming these issues in future studies would strengthen the quality of evidence on the psychometric properties of OB instruments.

See final published article at Sangmi P & Chang Dae L. (2023) Psychometric properties of self-reported instruments for occupational balance: A COSMIN-based systematic review, *Journal of Occupational Science*, DOI: 10.1080/14427591.2023.2243281

The unmet preconditions (i.e., sufficient internal consistency is conditional on sufficient structural validity) related to structural validity led to 'indeterminate (?)' overall ratings of the internal consistency of all five instruments. This result implies that the robustness of internal consistency needs to be evaluated not only on the magnitude of the coefficient of internal consistency but also on structural validity together with internal consistency. For development studies of new OB instruments, strengthening the robustness of structural validity in accordance with the recommended standards for self-reported instruments (Prinsen et al., 2018; Reeve et al., 2013) might be one way to strengthen the robustness of internal consistency. In this study, there were four studies which applied Rasch analysis to examine structural validity which was all rated as indeterminate. It was mainly due to the inconsistencies between the reported statistics and COSMIN standard. In addition, variation in reported statistics among the studies led to difficulties in comparison of the results of the studies. Following the Rasch reporting guidelines (Van de Winckel et al., 2022) would be helpful to solve the inconsistency with COSMIN standards and variation in reporting the results of Rasch analysis.

In addition to strengthening the quality of evidence and overall rating, it is recommended that researchers examine more diverse types of psychometric properties of OB instruments. Although the five psychometric properties were investigated in this study, three (i.e., structural validity, internal consistency, and reliability) were analysed within the COSMIN guidelines. Other psychometric properties (i.e., cross-cultural validity/measurement invariance and measurement error) were not analysable, primarily because they were not reported in the included studies. There was one included study (Håkansson et al., 2020) in which cross-cultural validity for the OBQ and OBQ11 were reported, however, the results were not analysable due to inconsistencies between reported estimates and standards of the COSMIN guidelines.

See final published article at Sangmi P & Chang Dae L. (2023) Psychometric properties of self-reported instruments for occupational balance: A COSMIN-based systematic review, *Journal of Occupational Science*, DOI: 10.1080/14427591.2023.2243281

Measurement error (e.g., the smallest detectable change and minimal important change) of the current OB instruments were not reported among the included studies. Additional studies to examine cross-cultural validity or measurement invariance and measurement error are encouraged to verify the research and clinical usefulness of OB instruments (van Kampen et al. 2013).

Implications for occupational science

The concept of OB is multidimensional (Dür et al., 2015; Wagman, Håkansson, & Björklund, 2012). The results of OB instruments which consist of the items of daily occupations (i.e., LBI, MAWNA, OB-Quest) provide the scores for subdomains. This allowed the instrument users to compare perceived OB between several dimensions of occupational engagement. It is notable that the subdomains of the three instruments are different from one another. This difference indicates that a variety of opinions of researchers for dimensions of OB or inconsistent characteristics of OB according to the study samples may exist. For this reason, further studies to clarify dimensions to be considered for measuring OB and to examine the appropriateness of applying the same dimensions to all individuals are recommended.

Achieving balance among various occupations is a dynamic process (Evans et al., 2014; Westhorp, 2003). Although all items within an OB instrument are equally rated using the same scale, the impact of an occupation in personal life may differ by individual. Likewise, balance across some specific occupational areas can have larger influence in perceived OB. Therefore, developing additional items or scoring calculation formulas to reflect the weighted value of each item can contribute to not underestimating or overestimating perceived OB. Considering the subjectivity is directly related to perceived OB (Backman & Forhan, 2017; Matuska, 2016),

See final published article at Sangmi P & Chang Dae L. (2023) Psychometric properties of self-reported instruments for occupational balance: A COSMIN-based systematic review, *Journal of Occupational Science*, DOI: 10.1080/14427591.2023.2243281

future studies on mechanisms of subjectivity change over time can contribute to understanding change in perceived OB. For example, life events which bring long lasting changes (ex. move to other job/city, getting married, retirement, etc.) could drop the levels of OB temporarily but OB may recover as the individual adapts to a new environment over time.

Limitations

The results of this study provide information on the limited number of psychometric properties of OB instruments compared to the number of psychometric properties that can be analysed using COSMIN guidelines. This was because the psychometric properties that did not align with the definitions of the COSMIN taxonomy (Mokkink et al., 2010) were excluded from the analysis. Additionally, some estimates were not ratable within the guidelines.

Some psychometric properties defined by the COSMIN taxonomy (Mokkink et al., 2010) were excluded from this study. For example, content validity was not considered in this study because researchers qualified to analyse the content validity (Terwee et al., 2018) did not participate in this study. Hypothesis testing for construct validity, responsiveness and criterion validity were not considered because the gold-standard instrument for measuring OB or consensus definition of the construct of OB does not exist. The unavailability of a gold-standard was also proved by the construct of OB in four instruments—LBI, MAWNA, OBQ, and OB-Quest—which were conceptualised and measured separately from each other (Håkansson et al., 2020). Fundamental studies on the concept of OB are also recommended to improve the accuracy of measuring it.

Findings of this study can be applied to self-reported measures for OB used in occupational science and occupational therapy. Other types of instruments for measuring OB

See final published article at Sangmi P & Chang Dae L. (2023) Psychometric properties of self-reported instruments for occupational balance: A COSMIN-based systematic review, *Journal of Occupational Science*, DOI: 10.1080/14427591.2023.2243281

(e.g., observer-rated measures and interview-based measures; Bejerholm et al., 2006; Eklund & Argentzell, 2016) or instruments developed in other disciplines (Davis et al., 2014) were not considered in this study.

Conclusions

OB instruments with a 'high' quality of evidence for sufficient psychometric properties are presently unavailable. The studies to strengthen the robustness of the psychometrics of the current instruments and to examine more diverse psychometrics of the current and future OB instruments are recommended. Also, further studies to better define the concept are needed in addition to deeper understanding of the characteristics of subjectivity to measure OB more accurately.

Funding details

This work was supported by the Ministry of Education (MOE) of the Republic of Korea and National Research Foundation (NRF) of Korea under BK21 FOUR (Fostering Outstanding Universities for Research) - Big data specialized education and research team for cognitive health and social integration of community-dwelling older adults [no grant number]; and the Paralyzed Veterans of America. The contents of this paper do not represent the views of Paralyzed Veterans of America.

Disclosure Statement

See final published article at Sangmi P & Chang Dae L. (2023) Psychometric properties of self-reported instruments for occupational balance: A COSMIN-based systematic review, *Journal of Occupational Science*, DOI: 10.1080/14427591.2023.2243281

The authors report there are no competing interests to declare.

Data Availability Statement

n/a

Acknowledgements

We thank Myung Ha Kim of Yonsei Wonju Medical Library for his assistance with the search data.

References

Backman, C. L. (2004). Occupational balance: Exploring the relationships among daily occupations and their influence on well-being. *Canadian Journal of Occupational Therapy*, *71*(4), 202-209. https://doi.org/10.1177/000841740407100404

Backman, C. L., & Forhan, M. (2017). Measuring time use and occupational balance. In M. Law, C. Baum, & W. Dunn (Eds.), *Measuring occupational performance: Supporting best practice in occupational therapy* (3rd ed., pp. 333-349). Slack.

Bejerholm, U., Hansson, L., & Eklund, M. (2006). Profiles of occupational engagement in people with schizophrenia (POES): The development of a new instrument based on time-use diaries. *British Journal of Occupational Therapy*, *69*(2), 58-68.

See final published article at Sangmi P & Chang Dae L. (2023) Psychometric properties of self-reported instruments for occupational balance: A COSMIN-based systematic review, *Journal of Occupational Science*, DOI: 10.1080/14427591.2023.2243281

https://doi.org/10.1177/030802260606900203

Christiansen, C. H., Baum, C. M., & Bass, J. D. (2015). Health, occupational performance, and occupational therapy. In C. H. Christiansen, C. M. Baum, & J. D. Bass (Eds.), *Occupational therapy: Performance, participation, and well-being* (4th ed., pp. 7-21). Slack.

Clouston, T. J. (2014). Whose occupational balance is it anyway? The challenge of neoliberal capitalism and work–life imbalance. *British Journal of Occupational Therapy*, *77*(10), 507-515. https://doi.org/10.4276/030802214X14122630932430

Clouston, T. J. (2015). *Challenging stress, burnout and rust-out: Finding balance in busy lives*. Jessica Kingsley Publishers.

Davis, R., Balkin, R. S., & Juhnke, G. A. (2014). Validation of the Juhnke–Balkin life balance inventory. *Measurement and Evaluation in Counseling and Development*, *47*(3), 181-198. https://doi.org/10.1177/0748175614531796

Dür, M., Steiner, G., Fialka-Moser, V., Kautzky-Willer, A., Dejaco, C., Prodinger, B., Stoffer, M. A., Binder, A., Smolen, J., & Stamm, T. A. (2014). Development of a new occupational balance-questionnaire: incorporating the perspectives of patients and healthy people in the design of a self-reported occupational balance outcome instrument. *Health and Quality of Life Outcomes*, *12*(1), 1-11. https://doi.org/10.1186/1477-7525-12-45

Dür, M., Unger, J., Stoffer, M., Drăgoi, R., Kautzky-Willer, A., Fialka-Moser, V., Smolen, J., & Stamm, T. (2015). Definitions of occupational balance and their coverage by instruments. *British Journal of Occupational Therapy*, *78*(1), 4-15. https://doi.org/10.1177/0308022614561235

Eakman, A. M. (2015). The meaningful activity wants and needs assessment: A perspective on

life balance. *Journal of Occupational Science*, *22*(2), 210-227. https://doi.org/10.1080/14427591.2013.769405

Eklund, M., & Argentzell, E. (2016). Perception of occupational balance by people with mental illness: A new methodology. *Scandinavian Journal of Occupational Therapy, 23*(4), 304-313. https://doi.org/10.3109/11038128.2016.1143529

Eklund, M., Orban, K., Argentzell, E., Bejerholm, U., Tjörnstrand, C., Erlandsson, L.K., & Håkansson, C. (2017). The linkage between patterns of daily occupations and occupational balance: applications within occupational science and occupational therapy practice. *Scandinavian Journal of Occupational Therapy*, *24*(1), 41-56. https://doi.org/10.1080/11038128.2016.1224271

Evans, K. L., Millsteed, J., Richmond, J. E., Falkmer, M., Falkmer, T., & Girdler, S. J. (2014). The complexity of role balance: Support for the Model of Juggling Occupations. *Scandinavian Journal of Occupational Therapy, 21*(5), 334-347. https://doi.org/10.3109/11038128.2014.902988

Fiore Jr, J. F., Figueiredo, S., Balvardi, S., Lee, L., Nauche, B., Landry, T., Mayo, N. E., & Feldman, L. S. (2018). How do we value postoperative recovery?: A systematic review of the measurement properties of patient-reported outcomes after abdominal surgery. *Annals of Surgery*, *267*(4), 656-669. https://doi.org/10.1097/SLA.0000000000002415

Gerber, L. H., & Furst, G. P. (1992). Validation of the NIH activity record: A quantitative measure of life activities. *Arthritis Care & Rheumatism, 5*(2), 81-86. https://doi.org/10.1002/art.1790050206

Harrison, J. L., Brownlow, C. L., Ireland, M. J., & Piovesana, A. M. (2022). Empathy measurement in autistic and nonautistic adults: A COSMIN systematic literature review.

Assessment, 29(2), 332-350. https://doi.org/10.1177/1073191120964564

Håkansson, C., Wagman, P., & Hagell, P. (2020). Construct validity of a revised version of the Occupational Balance Questionnaire. *Scandinavian Journal of Occupational Therapy*, *27*(6), 441-449. https://doi.org/10.1080/11038128.2019.1660801

Jewell, T., Gardner, T., Susi, K., Watchorn, K., Coopey, E., Simic, M., Fonagy, P., & Eisler, I. (2019). Attachment measures in middle childhood and adolescence: A systematic review of measurement properties. *Clinical Psychology Review, 68*, 71-82. https://doi.org/10.1016/j.cpr.2018.12.004

Jonsson, H., & Persson, D. (2006). Towards an experiential model of occupational balance: An alternative perspective on flow theory analysis. *Journal of Occupational Science*, *13*(1), 62-73. https://doi.org/10.1080/14427591.2006.9686571

Kielhofner, G. (2007). *Model of human occupation: Theory and application* (4th ed.). Lippincott Williams & Wilkins.

Kroenke, K., Monahan, P. O., & Kean, J. (2015). Pragmatic characteristics of patient-reported outcome measures are important for use in clinical practice. *Journal of Clinical Epidemiology*, *68*(9), 1085-1092. https://doi.org/10.1016/j.jclinepi.2015.03.023

Lee, C. D., Kim, M. Y., Lee, M. J., Kang, J., & Foster, E. R. (2023). The relationship between active, balanced participation and well-being in older adults in the United States: A time-use perspective. *Journal of Occupational Science*, *30*(2), 175-183. https://doi.org/10.1080/14427591.2020.1869584

Lexén, A., Kåhlin, I., Erlandsson, L.-K., & Håkansson, C. (2020). Occupational health among Swedish occupational therapists: A cross-sectional study. *International Journal of Environmental Research and Public Health, 17*(10), 3379.

See final published article at Sangmi P & Chang Dae L. (2023) Psychometric properties of self-reported instruments for occupational balance: A COSMIN-based systematic review, *Journal of Occupational Science*, DOI: 10.1080/14427591.2023.2243281

https://doi.org/10.3390/ijerph17103379

Little, B. R. (1983). Personal projects: A rationale and method for investigation. *Environment and Behavior, 15*(3), 273-309. https://doi.org/10.1177/0013916583153002

Matuska, K. (2012). Description and development of the Life Balance Inventory. *OTJR: Occupation, Participation and Health*, *32*(1), 220-228. https://doi.org/10.3928/15394492-20110610-0

Matuska, K. (2016). *Life balance: Science and stories of everyday living.* Balboa Press.

Mokkink, L. B., de Vet, H. C. W., Prinsen, C. A. C., Patrick, D. L., Alonso, J., Bouter, L. M., & Terwee, C. B. (2018). COSMIN Risk of Bias checklist for systematic reviews of Patient-Reported Outcome Measures. *Quality of Life Research*, *27*(5), 1171-1179. https://doi.org/10.1007/s11136-017-1765-4

Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., Bouter, L. M., & de Vet, H. C. (2010). The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *Journal of Clinical Epidemiology, 63*(7), 737-745. https://doi.org/10.1016/j.jclinepi.2010.02.006

Ohno, K., Tomori, K., Sawada, T., Seike, Y., Yaguchi, A., & Kobayashi, R. (2021). Measurement properties of the Canadian Occupational Performance Measure: A systematic review. *American Journal of Occupational Therapy, 75*(6), 7506205100. https://doi.org/10.5014/ajot.2021.041699

Ortiz-Rubio, A., Cabrera-Martos, I., Haro-Piedra, E., López-López, L., Rodríguez-Torres, J., Granados-Santiago, M., & Valenza, M. C. (2020). Exploring perceived occupational balance in women with fibromyalgia. A descriptive study. *Scandinavian Journal of*

Occupational Therapy, *29*(5), 395-402. https://doi.org/10.1080/11038128.2020.1865449

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., & Brennan, S. E. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, *372*(71). https://doi.org/10.1136/bmj.n71

Park, S., Lee, H. J., Jeon, B.J., Yoo, E.-Y., Kim, J.B., & Park, J.H. (2021). Effects of occupational balance on subjective health, quality of life, and health-related variables in community-dwelling older adults: A structural equation modeling approach. *Plos One*, *16*(2), e0246887. https://doi.org/10.1371/journal.pone.0246887

Pike, S., Lannin, N. A., Wales, K., & Cusick, A. (2018). A systematic review of the psychometric properties of the Action Research Arm Test in neurorehabilitation. *Australian Occupational Therapy Journal, 65*(5), 449-471. https://doi.org/10.1111/1440-1630.12527

Prinsen, C. A., Mokkink, L. B., Bouter, L. M., Alonso, J., Patrick, D. L., De Vet, H. C., & Terwee, C. B. (2018). COSMIN guideline for systematic reviews of patient-reported outcome measures. *Quality of Life Research*, *27*(5), 1147-1157. https://doi.org/10.1007/s11136-018-1798-3

Reeve, B. B., Wyrwich, K. W., Wu, A. W., Velikova, G., Terwee, C. B., Snyder, C. F., Schwartz, C., Revicki, D. A., Moinpour, C. M., & McLeod, L. D. (2013). ISOQOL recommends minimum standards for patient-reported outcome measures used in patient-centered outcomes and comparative effectiveness research. *Quality of Life Research, 22*(8), 1889-1905. https://doi.org/10.1007/s11136-012-0344-y

Riediger, M., Freund, A. M., & Baltes, P. B. (2005). Managing life through personal goals: Intergoal facilitation and intensity of goal pursuit in younger and older adulthood. *The*

Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 60(2), 84-91. https://doi.org/10.1093/geronb/60.2.p84

Rintala, D., Uttermohlen, D., Buck, E., Hanover, D., Alexander, J., & Norris-Baker, C. (1984). *Self-observation and report technique: Description and clinical applications* (A. Halpern & F. MJ, Eds.). Paul H. Brookes.

Sainburg, R. L., Liew, S.L., Frey, S. H., & Clark, F. (2017). Promoting translational research among movement science, occupational science, and occupational therapy. *Journal of Motor Behavior*, *49*(1), 1-7. https://doi.org/10.1080/00222895.2016.1271299

Shoman, Y., Marca, S., Bianchi, R., Godderis, L., Van der Molen, H., & Canu, I. G. (2021). Psychometric properties of burnout measures: A systematic review. *Epidemiology and Psychiatric Sciences, 30*. E8. https://doi.org/10.1017/S2045796020001134

Soh, S. L.H., Lane, J., Xu, T., Gleeson, N., & Tan, C. W. (2021). Falls efficacy instruments for community-dwelling older adults: a COSMIN-based systematic review. *BMC Geriatrics, 21*(1), 21. https://doi.org/10.1186/s12877-020-01960-7

Stamm, T., Lovelock, L., Stew, G., Nell, V., Smolen, J., Machold, K., Jonsson, H., & Sadlo, G. (2009). I have a disease but I am not ill: A narrative study of occupational balance in people with rheumatoid arthritis. *OTJR: Occupation, Participation and Health*, *29*(1), 32-39. https://doi.org/10.1177/153944920902900105

Terwee, C. B., Jansma, E. P., Riphagen, I. I., & de Vet, H. C. (2009). Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. *Quality of Life Research*, *18*(8), 1115-1123. https://doi.org/10.1007/s11136-009-9528-5

Terwee, C. B., Prinsen, C. A., Chiarotto, A., Westerman, M. J., Patrick, D. L., Alonso, J., Bouter,

L. M., De Vet, H. C., & Mokkink, L. B. (2018). COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. *Quality of Life Research*, *27*(5), 1159-1170. https://doi.org/10.1007/s11136-018-1829-0

Van de Winckel, A., Kozlowski, A. J., Johnston, M. V., Weaver, J., Grampurohit, N., Terhorst, L., ... & Mallinson, T. (2022). Reporting guideline for RULER: Rasch reporting guideline for rehabilitation research: Explanation and elaboration. *Archives of Physical Medicine and Rehabilitation, 103*(7), 1487-1498. https://doi.org/10.1016/j.apmr.2022.03.019

van Kampen, D. A., Willems, W. J., van Beers, L. W., Castelein, R. M., Scholtes, V. A., & Terwee, C. B. (2013). Determination and comparison of the smallest detectable change (SDC) and the minimal important change (MIC) of four-shoulder patient-reported outcome measures (PROMs). *Journal of Orthopaedic Surgery and Research, 8*(1), 1-9. https://doi.org/10.1186/1749-799X-8-40

Wagman, P., & Håkansson, C. (2014a). Introducing the occupational balance questionnaire (OBQ). *Scandinavian Journal of Occupational Therapy*, *21*(3), 227-231. https://doi.org/10.3109/11038128.2014.900571

Wagman, P., & Håkansson, C. (2014b). Exploring occupational balance in adults in Sweden. *Scandinavian Journal of Occupational Therapy, 21*(6), 415-420. https://doi.org/10.3109/11038128.2014.934917

Wagman, P., Håkansson, C., & Björklund, A. (2012). Occupational balance as used in occupational therapy: A concept analysis. *Scandinavian Journal of Occupational Therapy*, *19*(4), 322-327. https://doi.org/10.3109/11038128.2011.596219

Wagman, P., Håkansson, C., Jacobsson, C., Falkmer, T., & Björklund, A. (2012). What is considered important for life balance? Similarities and differences among some working

adults. *Scandinavian Journal of Occupational Therapy*, *19*(4), 377-384. https://doi.org/10.3109/11038128.2011.645552

Westhorp, P. (2003). Exploring balance as a concept in occupational science. *Journal of Occupational Science, 10*(2), 99-106.https://doi.org/10.1080/14427591.2003.9686516

Wilcock, A., Chelin, M., Hall, M., Hamley, N., Morrison, B., Scrivener, L., Townsend, M., & Treen, K. (1997). The relationship between occupational balance and health: A pilot study. *Occupational Therapy International*, *4*(1), 17-30. https://doi.org/10.1002/oti.45

Wilcock, A. A. (1991). Occupational science. *British Journal of Occupational Therapy*, *54*(8), 297-300. https://doi.org/10.1177/030802269105400807

Wilcock, A. A. (2006). *An occupational perspective of health* (2nd ed.). Slack.

Wilcock, A. A., & Hocking, C. (2015). *An occupational perspective of health* (3rd ed.). Slack.

Yuen, H. K., & Austin, S. L. (2014). Systematic review of studies on measurement properties of instruments for adults published in the American Journal of Occupational Therapy, 2009–2013. *American Journal of Occupational Therapy, 68*(3), e97-e106. https://doi.org/10.5014/ajot.2014.011171

See final published article at Sangmi P & Chang Dae L. (2023) Psychometric properties of self-reported instruments for occupational balance: A COSMIN-based systematic review, *Journal of Occupational Science*, DOI: 10.1080/14427591.2023.2243281

Appendix 1.

Cochrane Strategy

#	Searches
#1	("Occupational Balance-Questionnaire" OR "Occupational balance Questionnaire" OR "Life Balance Inventory" OR "Meaningful Activity Wants and Needs Assessment" OR "OB-Quest" OR OBQ OR LBI OR MAWNA):ti,ab,kw" (Word variations have been searched)
#2	limit to Year first published (2010-present)

Web of Science search strategy

#	Searches
#1	TS= ("Occupational Balance-Questionnaire" OR "Occupational balance Questionnaire" OR "Life Balance Inventory" OR "Meaningful Activity Wants and Needs Assessment" OR "OB-Quest" OR OBQ OR LBI OR MAWNA) AND Language: (English)
#2	Editions=SCI-EXPANDED, SSCI, A&HCI, ESCI Publication Date=2010-2021

Pubmed (https://pubmed.ncbi.nlm.nih.gov/)

#	Searches
1	"Occupational Balance-Questionnaire"[Text Word] OR "Occupational balance Questionnaire"[Text Word] OR "Life Balance Inventory"[Text Word] OR "Meaningful Activity Wants and Needs Assessment"[Text Word] OR "OB-Quest"[Text Word] OR OBQ[Text Word] OR LBI[Text Word] OR MAWNA[Text Word]

See final published article at Sangmi P & Chang Dae L. (2023) Psychometric properties of self-reported instruments for occupational balance: A COSMIN-based systematic review, *Journal of Occupational Science*, DOI: 10.1080/14427591.2023.2243281

(instrumentation[sh] OR methods[sh] OR "Validation Study"[pt] OR "Comparative Study"[pt] OR "psychometrics"[MeSH] OR psychometr*[tiab] OR clinimetr*[tw] OR clinometr*[tw] OR "Outcome Assessment, Health Care"[Mesh] OR "outcome assessment"[tiab] OR "outcome measure*"[tw] OR "observer variation"[MeSH] OR "observer variation"[tiab] OR "Health Status Indicators"[Mesh] OR "reproducibility of results"[MeSH] OR reproducib*[tiab] OR "discriminant analysis"[MeSH] OR reliab*[tiab] OR unreliab*[tiab] OR valid*[tiab] OR "coefficient of variation"[tiab] OR coefficient[tiab] OR homogeneity[tiab] OR homogeneous[tiab] OR "internal consistency"[tiab] OR (cronbach*[tiab] AND (alpha[tiab] OR alphas[tiab])) OR (item[tiab] AND (correlation*[tiab] OR selection*[tiab] OR reduction*[tiab])) OR agreement[tw] OR precision[tw] OR imprecision[tw] OR "precise values"[tw] OR test-retest[tiab] OR (test[tiab] AND retest[tiab]) OR (reliab*[tiab] AND (test[tiab] OR retest[tiab])) OR stability[tiab] OR interrater[tiab] OR inter-rater[tiab] OR intrarater[tiab] OR intra-rater[tiab] OR intertester[tiab] OR inter-tester[tiab] OR intratester[tiab] OR intra-tester[tiab] OR interobserver[tiab] OR inter-observer[tiab] OR intraobserver[tiab] OR intra-observer[tiab] OR intertechnician[tiab] OR inter-technician[tiab] OR intratechnician[tiab] OR intra-technician[tiab] OR interexaminer[tiab] OR inter-examiner[tiab] OR intraexaminer[tiab] OR intra-examiner[tiab] OR interassay[tiab] OR inter-assay[tiab] OR intraassay[tiab] OR intra-assay[tiab] OR interindividual[tiab] OR inter-individual[tiab] OR intraindividual[tiab] OR intra-individual[tiab] OR interparticipant[tiab] OR inter-participant[tiab] OR intraparticipant[tiab] OR intra-participant[tiab] OR kappa[tiab] OR kappa's[tiab] OR kappas[tiab] OR repeatab*[tw] OR ((replicab*[tw] OR repeated[tw]) AND (measure[tw] OR measures[tw] OR findings[tw] OR result[tw] OR results[tw] OR test[tw] OR tests[tw])) OR generaliza*[tiab] OR generalisa*[tiab] OR concordance[tiab] OR (intraclass[tiab] AND correlation*[tiab]) OR discriminative[tiab] OR "known group"[tiab] OR "factor analysis"[tiab] OR "factor analyses"[tiab] OR "factor structure"[tiab] OR "factor structures"[tiab] OR dimension*[tiab] OR subscale*[tiab] OR (multitrait[tiab] AND scaling[tiab] AND (analysis[tiab] OR analyses[tiab])) OR "item discriminant"[tiab] OR "interscale correlation*"[tiab] OR error[tiab] OR errors[tiab] OR "individual variability"[tiab] OR "interval variability"[tiab] OR "rate variability"[tiab] OR (variability[tiab] AND (analysis[tiab] OR values[tiab])) OR (uncertainty[tiab] AND (measurement[tiab] OR measuring[tiab])) OR "standard error of measurement"[tiab] OR sensitiv*[tiab] OR responsive*[tiab] OR (limit[tiab] AND detection[tiab]) OR "minimal detectable concentration"[tiab] OR interpretab*[tiab] OR ((minimal[tiab] OR minimally[tiab] OR clinical[tiab] OR clinically[tiab]) AND (important[tiab] OR significant[tiab] OR detectable[tiab]) AND (change[tiab] OR difference[tiab])) OR (small*[tiab] AND (real[tiab] OR detectable[tiab]) AND (change[tiab] OR difference[tiab])) OR "meaningful change"[tiab] OR "ceiling effect"[tiab] OR "floor effect"[tiab] OR "Item response model"[tiab] OR IRT[tiab] OR Rasch[tiab] OR "Differential item functioning"[tiab] OR DIF[tiab] OR "computer adaptive testing"[tiab] OR "item bank"[tiab] OR "cross-cultural equivalence"[tiab])

2

3 #1 AND #2

("delphi-technique"[ti] OR cross-sectional[ti] OR "address"[Publication Type] OR "biography"[Publication Type] OR "case reports"[Publication Type] OR "comment"[Publication Type] OR "directory"[Publication Type] OR "editorial"[Publication Type] OR "festschrift"[Publication Type] OR "interview"[Publication Type] OR "lecture"[Publication Type] OR "legal case"[Publication Type] OR "legislation"[Publication Type] OR "letter"[Publication Type] OR "news"[Publication Type] OR "newspaper article"[Publication Type] OR "patient education handout"[Publication Type] OR "popular work"[Publication Type] OR "congress"[Publication Type] OR "consensus development conference"[Publication Type] OR "consensus development conference, nih"[Publication Type] OR "practice guideline"[Publication Type]) NOT ("animals"[MeSH Terms] NOT "humans"[MeSH Terms])

4

5 #3 NOT #4

See final published article at Sangmi P & Chang Dae L. (2023) Psychometric properties of self-reported instruments for occupational balance: A COSMIN-based systematic review, *Journal of Occupational Science*, DOI: 10.1080/14427591.2023.2243281

6	("2010/01/01"[Date - Publication] : "3000"[Date - Publication])
7	#5 AND #6
8	"English"[Language]
9	#7 AND #8

Psycinfo(Proquest) search strategy

#	**Searches**
1	"Occupational Balance-Questionnaire" OR "Occupational balance Questionnaire" OR "Life Balance Inventory" OR "Meaningful Activity Wants and Needs Assessment" OR "OB-Quest" OR OBQ OR LBI OR MAWNA
2	Applied filters: Publication date (2010-01-01 - 2021), Language (English)

Appendix 2.

Characteristics of the included OB instruments

Instrument	Developer/s (year of publication)	County/ Language	Total items/ Domains (items)	Response options	Range of scores (scoring)	Interpretation
LBI	Matuska (2012)	USA/ English	53/ Health(6), Relationship(10), Identity(16), Challenge(19)	3-point [a]	1-3 (Take average for each domain) [b]	A higher score indicates better occupational balance
MAWNA	Eakman	USA/	21/Competence and	4-point	0-63	A higher score

	(2015)	English	goal achievement (9), Pleasure and enjoyment (7), Social connectedness (5)		(Sum up)	indicates higher unmet needs toward meaningful experiences
OBQ	Wagman & Håkansson (2014b)	Sweden/ Swedish	13/ -	6-point	0-65 (Sum up) / 0-5 (item) [c]	A higher score indicates better occupational balance
OBQ11	Håkansson et al. (2020)	Sweden/ Swedish	11/ -	4-point	0-33 (Sum up) / 0-3 (item) [c]	A higher score indicates better occupational balance
OB-Quest	Dür et al. (2014)	Austria/ German	10/ challenging and relaxing activities (2), Activities with acknowledgement (1), Health impact on activities (1), Stress in activities (1), Rest and sleep (2), Variety of activities (1), and Adaptation of	3-point	1-3 /2-6 (Sum up for each domain) [d]	A lower score indicates better occupational balance

See final published article at Sangmi P & Chang Dae L. (2023) Psychometric properties of self-reported instruments for occupational balance: A COSMIN-based systematic review, *Journal of Occupational Science*, DOI: 10.1080/14427591.2023.2243281

activities (2)

Note. LBI=Life Balance Inventory; MAWNA=Meaningful Activity Wants and Needs Assessment; OB-Quest=Occupational Balance Questionnaire; OBQ=Occupational Balance Questionnaire.

[a] (1) always less/more than I want, (2) sometimes less/more than I want, and (3) about right for me.

[b] Respondents are asked to score the items which indicate the activities they have participated in within last four weeks or they want to participated in.

[c] Instrument users may use not only summed score but also each item score for analysis.

[d] Range of scores for a domain composed of a single item is 1 - 3. Range of scores for a domain composed of two items is 2 - 6.

See final published article at Sangmi P & Chang Dae L. (2023) Psychometric properties of self-reported instruments for occupational balance: A COSMIN-based systematic review, *Journal of Occupational Science*, DOI: 10.1080/14427591.2023.2243281