

Publication status: This preprint has not been published elsewhere.

Work-Related Quality of Life Scale (WRQoL-2): Differential Item Functioning Analysis

Pedro Afonso Cortez, Heila Magali da Silva Veiga, Helenides Mendonça, Elaine Rabelo Neiva,
Darren Van Laar

<https://doi.org/10.1590/1982-4327e3628>

Submitted on: 2026-10-20

Posted on: 2026-10-20 (version 1)

(YYYY-MM-DD)

Paidéia

2026, Vol. 36, e3628. <https://doi.org/10.1590/1982-4327e3628>

ISSN 1982-4327 (online version)

Thematic Dossier

Work-Related Quality of Life Scale (WRQoL-2): Differential Item Functioning Analysis

Pedro Afonso Cortez¹ – <https://orcid.org/0000-0003-0107-2033>

Heila Magali da Silva Veiga¹ – <https://orcid.org/0000-0002-7429-8124>

Helenides Mendonça² – <https://orcid.org/0000-0003-3990-5432>

Elaine Rabelo Neiva³ – <https://orcid.org/0000-0003-0503-3234>

Darren Van Laar⁴ – <https://orcid.org/0000-0002-4424-0124>

Abstract: This study examined reliability and validity evidence for the Work-Related Quality of Life Scale (WRQoL-2) in Brazil using Item Response Theory (IRT). The sample comprised 209 adults (168 with and 41 without a mental health diagnosis) to assess Differential Item Functioning (DIF). Person reliability was .930. Although the likelihood ratio test favoured the Partial Credit Model (PCM), AIC, BIC, and CAIC favoured the more parsimonious Rating Scale Model (RSM). Five items showed elevated infit/outfit values, and 13 of 31 items exhibited disordered thresholds. Items 9 and 10 showed significant DIF after Holm adjustment ($p = .036$). Given the small non-diagnosed group, lack of DIF in other items does not prove absence of bias. In conclusion, these two items require revision prior to subgroup comparisons.

Keywords: social welfare, mental health, behavior rating scale, self-testing, psychometrics

¹*Universidade Federal de Uberlândia, Uberlândia-MG, Brazil.*

²*Pontifícia Universidade Católica de Goiás, Goiânia-GO, Brazil.*

³*Universidade de Brasília, Brasília-DF, Brazil.*

⁴*University of Portsmouth, Portsmouth, United Kingdom.*

Correspondence address: Pedro Afonso Cortez. Universidade Federal de Uberlândia, Instituto de Psicologia, Av. Pará, 1720, Bloco 2C, Campus Umuarama, CEP: 38405-240, Uberlândia - Minas Gerais, Brazil.
pedro.cortez@ufu.br

Escala de Qualidade de Vida no Trabalho (WRQoL-2): Análise de Funcionamento

Diferencial dos Itens

Resumo: Este estudo investigou evidências de confiabilidade e validade da Escala de Qualidade de Vida no Trabalho (WRQoL-2) no Brasil, via Teoria de Resposta ao Item (TRI). A amostra contou com 209 adultos (168 com e 41 sem diagnóstico de saúde mental) para analisar o Funcionamento Diferencial do Item (DIF). A confiabilidade das pessoas foi de 0,930. O teste de razão de verossimilhança favoreceu o Modelo de Crédito Parcial (PCM), porém os critérios AIC, BIC e CAIC favoreceram o Modelo de Escala de Classificação (RSM), mais parcimonioso. Cinco itens apresentaram *infit/outfit* elevados e 13 dos 31 itens exibiram limiares desordenados. Os itens 9 e 10 apresentaram DIF significativo ($p = 0,036$). Devido ao tamanho reduzido do grupo sem diagnóstico, a ausência de DIF nos demais itens não garante isenção de viés. Conclui-se que esses dois itens exigem revisão antes de comparações entre subgrupos.

Palavras-chave: bem-estar social; saúde mental; escala de avaliação comportamental; autoavaliação; psicometria

Escala de Calidad de Vida en el Trabajo (WRQoL-2): Análisis de Funcionamiento

Diferencial de los Ítems

Resumen: Este estudio analizó evidencias de confiabilidad y validez de la Escala de Calidad de Vida en el Trabajo (WRQoL-2) en Brasil mediante la Teoría de Respuesta al Ítem (TRI). La muestra incluyó 209 adultos (168 con y 41 sin diagnóstico de salud mental) para evaluar el Funcionamiento Diferencial de los Ítems (DIF). La confiabilidad de las personas fue de .930. La prueba de razón de verosimilitud favoreció el Modelo de Crédito Parcial (PCM), pero los criterios AIC, BIC y CAIC favorecieron el Modelo de Escala de Clasificación (RSM), más

Cortez, P. A. et al., (2026). WRQoL-2: Differential Item Functioning Analysis

parsimonioso. Cinco ítems presentaron infit/outfit elevados y 13 de los 31 ítems mostraron umbrales desordenados. Los ítems 9 y 10 presentaron DIF significativo ($p = .036$). Dado el tamaño reducido del grupo sin diagnóstico, la ausencia de DIF en otros ítems no demuestra ausencia de sesgo. Se concluye que ambos ítems requieren revisión antes de realizar comparaciones entre subgrupos.

Palabras clave: bienestar social, salud mental, escala de evaluación de la conducta, autoevaluación, psicometría

Assessing the quality of work life is crucial for understanding and improving employee well-being, which has a direct impact on organizational performance and effectiveness (Garzaro et al., 2020; Sezer et al., 2024). One of the key constructs in this area is the Work-Related Quality of Life (WRQoL) framework, which provides a comprehensive evaluation of various aspects of employees' work experiences (Garzaro et al., 2020).

The WRQoL framework encompasses several important dimensions, including job satisfaction, work-life balance, and stress management (Sabino et al., 2024; Silarova et al., 2022). These dimensions shape individual well-being and are also associated with productivity, employee engagement, and retention within the workplace (Van Vo et al., 2024). Evidence obtained in a propensity-score-matched sample of nurses further indicates that lower work-related quality of life is associated with a higher prevalence of depressive symptoms (Li, Chen, et al., 2023), which supports the relevance of examining mental health when work-related quality of life is measured.

The WRQoL scale, originally developed to provide a broad evaluation of work-related factors, has undergone refinement to improve its precision and reliability, resulting in the creation

of the WRQoL-2. This updated version incorporates more targeted items and improved psychometric properties. The WRQoL-2 measures seven key dimensions: Control at Work, Employee Engagement, General Well-Being, Home-Work Interface, Job Career Satisfaction, Stress at Work, and Working Conditions (Easton & Van Laar, 2013, 2018).

The psychometric quality of the WRQoL-2 has been supported in several samples (Silarova et al., 2022). In Portugal, the WRQoL-2 demonstrated a six-factor structure, with adequate reliability, convergent and discriminant validity, and invariance across gender and sector, as well as low-intensity associations with organizational performance (Sabino et al., 2024). The revised structure has also been examined in China (Li, Wang, et al., 2022). In Brazil, a confirmatory factor analysis conducted with the same sample analysed here supported the seven correlated dimensions of the instrument, with excellent fit ($CFI = TLI = .99$; $RMSEA = .01$, 95% CI [.01, .02]) and high total-score reliability ($\omega = .950$; $\alpha = .938$); correlations with a mental health battery supported criterion validity (Cortez et al., 2025).

While the WRQoL-2 has demonstrated reliability and validity in different cultural and professional settings, applying it in Brazil introduces specific challenges. Brazilian work culture combines economic instability, heterogeneous labour market conditions, and workplace norms shaped by local cultural values, which produce distinct work-life dynamics and make it necessary to examine how the WRQoL-2 functions in this context (Sinval et al., 2020).

A critical component of this validation process involves analysing item fit statistics and Differential Item Functioning (DIF) across subgroups, particularly those defined by mental health status (Hagquist & Andrich, 2017). Infit, or information-weighted fit, evaluates whether respondents whose trait level is close to the item location provide expected responses, and it is therefore informative about item performance in the region of the continuum where the item is most precise (Jones et al., 2023). Outfit, or outlier-sensitive fit, is an unweighted mean-square

Cortez, P. A. et al., (2026). WRQoL-2: Differential Item Functioning Analysis

statistic and is consequently dominated by responses located far from the item, which makes it comparatively more sensitive to a small number of aberrant or careless answers than to systematic misfit. Because the two statistics weight residuals differently, an item with acceptable infit but elevated outfit typically signals isolated unexpected responses, whereas elevated values in both statistics indicate that the item does not discriminate as the model expects along the whole continuum (Müller, 2020). Values close to 1.0 in both statistics suggest that an item functions consistently within the scale (Hrnjicic & Alihodžic, 2024).

DIF analysis indicates whether specific items favour one subgroup over another among respondents at the same level of the latent trait, and identifying significant DIF is essential to ensure that the scale measures the construct fairly across populations (Russell et al., 2021). Employees with mental health conditions may respond differently to certain items because of the challenges they face, such as stigma, heightened stress, or reduced autonomy at work, which may compromise the fairness of the data and the validity of the scale if left unaddressed (Olteanu et al., 2019).

For example, items such as “I feel able to voice opinions and influence changes in my area of work” or “I am involved in decisions that directly affect members of the public” may be interpreted differently by employees with mental health challenges, possibly reflecting perceived barriers such as limited psychological safety or organizational support (Østerud, 2023). Focusing on mental health subgroups is relevant because these conditions may heighten sensitivity to work-related factors such as stress, job design, communication, and managerial support (Hess, 2024; Wang et al., 2023).

The primary objective of this study was to examine the item-level functioning of the WRQoL-2 in a Brazilian sample using Item Response Theory (IRT), and to identify items with

potential bias through Differential Item Functioning (DIF) analysis across subgroups defined by mental health status. Three expectations guided the analyses. First, given that a confirmatory factor analysis conducted with this sample the instrument with a highly reliable total score (Cortez et al., 2025), the item set was expected to be unidimensional to a degree sufficient for a diagnostic Rasch-family calibration. Second, because the response categories of the WRQoL-2 were not developed for Brazilian respondents, some items were expected to display disordered thresholds and elevated fit statistics. Third, DIF was expected to concentrate in items referring to voice, participation, and organizational communication rather than to be distributed across the instrument. Because the subgroups differ considerably in size, the analyses are best understood as a diagnostic screening of item behaviour: they indicate which items warrant refinement and which patterns require confirmation in larger and more balanced samples. Within this scope, the study provides the first item-level evidence for the Brazilian WRQoL-2 and a set of specific, testable targets for its refinement.

Method

Participants

The study involved 209 economically active workers who had been in their current job for at least six months. Participants were categorized into two groups: 168 individuals identified as mental health patients and 41 as non-mental health patients, based on pre-consented access to their personal records. The participants' average age was 37.1 years ($SD = 13.0$), with an age range of 18 to 68 years. The sample included 161 white Brazilians (77.0%) and 48 non-white Brazilians (23.0%), 154 women (73.7%) and 55 men (26.3%); 183 participants (87.6%) had completed a degree and 26 (12.4%) had not. The mental health group corresponded to 80.4% of the sample. Inclusion criteria required being an economically active worker aged 18 or older with

Cortez, P. A. et al., (2026). WRQoL-2: Differential Item Functioning Analysis

at least six months in the current job, while exclusion criteria applied to individuals without formal employment or with incomplete consented records.

Instruments

Work-Related Quality of Life Scale (WRQoL-2, 1st Section): The WRQoL-2 is a self-reported instrument designed to evaluate employees' quality of working life. It measures seven key dimensions that are summarized in a general score of work-related quality of life: Control at Work (CAW), Employee Engagement (EEN), General Well-Being (GWB), Home-Work Interface (HWI), Job Career Satisfaction (JCS), Stress at Work (SAW), and Working Conditions (WCS). Responses are recorded on a five-point Likert scale ranging from '1 = Strongly Disagree' to '5 = Strongly Agree', and higher scores reflect a more positive perception of quality of working life. Content validity has been confirmed through literature reviews and expert panels, and a user manual provides instructions for administration, scoring, and interpretation (Easton & Van Laar, 2018). Of the 32 items composing the first section, 31 items constitute the seven dimensions and were retained for calibration; the remaining item is a global indicator of overall quality of working life that does not load on any dimension and was therefore excluded from the analyses. Psychometric evidence for the Brazilian version was obtained with the same sample analysed here: confirmatory factor analysis supported the seven correlated dimensions with excellent fit ($\chi^2(427) = 403.31$, $p = .78$; CFI = TLI = .99; RMSEA = .01, 95% CI [.01, .02]; SRMR = .06), item loadings above .40, reliability between $\omega = .790$ and $\omega = .912$ by dimension, and criterion validity against the Mental Health Assessment Battery - Pandemic Version (Peres & Cortez, 2024). That analysis modelled seven correlated factors and did not test a second-order or bifactor structure. The decision to calibrate the 31 items along a single continuum in the present study therefore rests on the coherence of the total score rather than on a

formally estimated general factor: the item pool shows a high degree of common variance ($KMO = .94$), total-score reliability is high ($\omega = .950$; $\alpha = .938$), and each dimension correlates strongly with the general score (absolute r between $.486$ and $.927$). These properties support treating the instrument as approximately unidimensional for the diagnostic purposes of this study. The extent to which this approximation holds is examined below through the test of local independence and is revisited among the limitations.

Sociodemographic and Functional Section (WRQoL-2 Final Section): A structured block of questions at the end of the instrument characterized the participants, recording age, race/ethnicity, gender, highest level of education, length of time in the current role (minimum of six months as an inclusion criterion), and employment status. With participants' prior consent, access to clinical records was requested in order to stratify the sample into mental health patients ($n = 168$) and non-mental health patients ($n = 41$).

Procedures

Data Collection. The study complied with ethical guidelines as approved by the Brazilian Ethics Committee. Prior to participation, all participants provided informed consent. The study was registered under the number CAAE <<ocultado para garantir blind review>>. Ethical protocols were followed to ensure participant confidentiality and voluntary involvement throughout data collection. Data were gathered through an online survey hosted on SurveyMonkey, where participants completed a structured questionnaire and could offer access through an appendix detailing medical diagnoses. Recruitment was conducted using a snowball sampling method, focusing on public organizational groups on social networks, with survey completion taking approximately 20 to 30 minutes.

Data Analysis. The data analysis was conducted using a set of statistical procedures to evaluate the reliability, validity, and item functioning of the WRQoL-2. Descriptive statistics

Cortez, P. A. et al., (2026). WRQoL-2: Differential Item Functioning Analysis

were first calculated for sociodemographic variables, including means, standard deviations, and ranges for continuous data and frequencies and percentages for categorical data. The analyses were performed using Jamovi version 2.3.28, with the snowIRT add-on for R 4.4.2 integrated to support IRT modelling. The Partial Credit Model (PCM) and the Rating Scale Model (RSM) were both applied to compare item functioning, as these models are well-suited for evaluating polytomous response data (Wind, 2023).

The two models differ specifically in how they treat threshold parameters. In the RSM, the distances between adjacent category thresholds are constrained to be identical across items, so each item is described by a single location parameter plus a set of threshold distances shared by the whole instrument. In the PCM, threshold distances are estimated freely for every item, so each item has its own set of thresholds. The PCM is therefore more flexible and can represent items whose response categories function differently from one another, at the cost of estimating substantially more parameters, whereas the RSM is more parsimonious but assumes that all items share the same category structure (Raykov & Pusic, 2025). Comparing the two models is informative because a preference for the RSM indicates that a common category structure is sufficient, whereas a preference for the PCM indicates that items require individual category structures.

Model fit was assessed using a range of indices, including log-likelihood, deviance, Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and Consistent Akaike Information Criterion (CAIC), and a likelihood ratio test was conducted to compare the models. Because the information criteria penalize model complexity while the likelihood ratio test does not, the two procedures can disagree; when this occurs, the information criteria provide the more conservative basis for model selection (Wind, 2023). Item-level fit was analysed using

infit and outfit statistics to assess response consistency, with values close to 1.0 indicating that items were functioning as expected within the model (Müller, 2020). In addition, threshold ordering was inspected for every item: thresholds are expected to increase monotonically across adjacent categories, and a decrease indicates that the corresponding intermediate category is never the most probable response at any point of the latent continuum, which constitutes category malfunctioning (Wind, 2023).

Person reliability was computed as the Rasch-model analogue of internal consistency, expressing the proportion of observed variance in person measures that is not attributable to measurement error; values above .80 indicate that the item set separates respondents reliably along the latent continuum (Lang et al., 2019). The assumption of local independence was evaluated with the mean of the absolute values of the centred Q3 statistic (MADaQ3), which summarizes the residual correlations among all item pairs after the Rasch model has been fitted. The associated p-value is obtained by comparing the observed MADaQ3 with a parametric bootstrap null distribution; a significant p-value indicates that residual correlations exceed the values expected under the model and therefore that local independence is not fully supported (Christensen et al., 2017).

To explore potential response biases, DIF analysis was performed to identify items that may behave differently across subgroups, comparing participants with and without mental health diagnoses (Hagquist & Andrich, 2017). Because 31 items were tested simultaneously, the p-values of the likelihood ratio DIF tests were adjusted with the Holm step-down procedure, which controls the family-wise error rate without assuming independence among tests and is less conservative than the Bonferroni correction (Holm, 1979). Two features of the design should be noted in advance. First, the subgroups are strongly unbalanced ($n = 168$ versus $n = 41$). Likelihood ratio DIF tests for polytomous items lose considerable power when one group is

Cortez, P. A. et al., (2026). WRQoL-2: Differential Item Functioning Analysis

small, so non-significant results are ambiguous: they are compatible both with the absence of DIF and with insufficient power to detect it (Belzak, 2020). Second, the likelihood ratio test is an omnibus test and does not indicate the direction of the difference between groups. The DIF results are therefore reported as a screening procedure for items requiring revision, and not as evidence of measurement invariance.

Ethics Committee Approval Statement

This study was conducted in accordance with the Brazilian National Health Council Resolution 510/2016. The study protocol was reviewed and approved by the Research Ethics Committee under the Certificate of Presentation of Ethical Review (CAAE: 37798820.8.0000.5508). Participation was entirely voluntary, and respondents were free to withdraw from the survey at any time without penalty or justification. Before accessing the online questionnaire, all participants were presented with an electronic Informed Consent Form detailing the study objectives, procedures, potential risks, and data handling protocols. Access to the questionnaire was granted only after participants confirmed their consent electronically. To guarantee anonymity and confidentiality, all collected data were completely de-identified, and no personally identifiable information (such as names, email addresses, or IP addresses) was recorded or stored.

Results

Person reliability was .930, indicating that the item set separated respondents reliably along the latent continuum. The mean of the absolute values of the centred Q3 statistic (MADaQ3) was .182, with a significant p-value ($< .001$) after Holm adjustment, indicating that

the residual correlations among item pairs exceeded the values expected under the Rasch model. The assumption of local independence was therefore not fully supported in this sample. This result is consistent with the organization of the WRQoL-2 into seven correlated dimensions and implies that the person reliability reported above should be read as an upper bound (Christensen et al., 2017). Two models were assessed: the Partial Credit Model (PCM) and the Rating Scale Model (RSM) for polytomous responses.

The PCM presented a log-likelihood of -8900 , a deviance of 17801 , and an AIC of 18051 , while the RSM had a log-likelihood of -8975 , a deviance of 17949 , and an AIC of 18019 . The BIC for the PCM was 18468 compared to 18136 for the RSM, with CAIC values of 18593 for the PCM and 18171 for the RSM. The PCM utilized 125 parameters, while the RSM employed 35 parameters, both models covering 209 participants. The likelihood ratio test yielded a chi-square (χ^2 , $df = 90$) = 149 with a significant p-value ($< .001$).

These figures are internally consistent but point in opposite directions. The likelihood ratio test favours the PCM, because its additional 90 parameters produce a statistically detectable improvement in likelihood. All three information criteria, which penalize this additional complexity, favour the RSM, and the differences are substantial: $\Delta AIC = 32$, $\Delta BIC = 332$, and $\Delta CAIC = 422$, all in favour of the more parsimonious model. Given that the PCM estimates 125 parameters from 209 respondents, a parameter-to-person ratio of 0.60 , the gain in likelihood is not accompanied by evidence that item-specific thresholds are estimated stably. Accordingly, this study does not claim that the PCM fits these data better than the RSM. The RSM is the model preferred on parsimony-penalized grounds, and the PCM is retained in the analyses that follow only as a diagnostic device, because its item-specific thresholds are what make it possible to screen individual items for category malfunctioning and misfit (Wind, 2023). The item-level analysis of the PCM is presented in Table 1.

Cortez, P. A. et al., (2026). WRQoL-2: Differential Item Functioning Analysis

Table 1*Delta-Tau Parameterization of the Partial Credit IRT Model*

Item	Item wording (Brazilian Portuguese)	Threshold parameters				Infit	Outfit
		1	2	3	4		
1	Tenho um conjunto de objetivos e metas claras que me permitem fazer meu trabalho	-.63	-.44	-.27	1.34	.76	.78
2	Tenho a oportunidade de usar minhas habilidades no trabalho	-.67	-.46	.07	1.05	.69	.67
3	Quando eu faço um bom trabalho, sou reconhecido pela chefia imediata	-.61	-.32	.01	.91	.69	.65
4	Sinto-me encorajado a desenvolver novas habilidades	-.64	-.49	.02	1.11	.74	.72
5	Estou satisfeito com as oportunidades de carreira disponíveis onde trabalho	-.79	.20	-.28	.87	.76	.75
6	Estou satisfeito com o treinamento que recebo para desempenhar meu trabalho atual	-.65	.01	-.10	.75	.92	.89
7	Sinto-me capaz de dar opiniões e influenciar mudanças em meu ambiente de trabalho	-.79	-.02	.25	.55	2.36	2.73
8	Eu estou envolvido nas tomadas de decisões que afetam o meu trabalho	-.56	-.44	.33	.67	.76	.74
9	Estou envolvido em decisões que afetam os clientes e o público da empresa que trabalho	-.39	.19	.16	.05	3.10	4.20
10	Tenho oportunidades suficientes para questionar minha chefia sobre mudanças no trabalho	-.53	-.60	.15	.98	.83	.86
11	Sinto-me bem neste momento	-.90	-.12	.09	.93	.85	.84
12 (R)	Recentemente, tenho me sentido triste e deprimido	-.45	-.37	.31	.51	.81	.80
13	Estou satisfeito com a minha vida	-.72	-.51	.41	.81	.70	.66
14	Em muitos aspectos minha vida é próxima do ideal	-.61	-.38	.33	.66	.75	.73
15	Geralmente, as coisas dão certo para mim	-.83	-.37	.09	1.11	.67	.67
16	De maneira geral, tenho me sentido feliz com as coisas que me aconteceram até aqui na minha vida e no trabalho	-.48	-.43	.20	.71	.86	.82
17	Meu empregador provê instalações adequadas e flexibilidade para eu equilibrar minha vida entre família e trabalho	-1.21	-.38	-.04	1.60	.61	.63
18	Minha carga horária atual se adequa bem às minhas rotinas de vida pessoal	-.72	-.06	-.10	.88	.74	.73
19	Minha chefia imediata possibilita uma jornada de trabalho flexível	-1.05	.37	.14	.54	2.66	3.15
20	Sou capaz de equilibrar de forma saudável a vida profissional e pessoal	-.68	-.28	.16	.80	.71	.68
21 (R)	Frequentemente sinto-me sob pressão no trabalho	-.50	-.76	.29	.98	.72	.77
22 (R)	Eu me sinto muito estressado no trabalho	-.69	-.43	.19	.93	.58	.57
23 (R)	Tenho prazos impossíveis de serem cumpridos	-.34	-.39	.32	.40	1.01	1.00
24 (R)	Sou pressionado a trabalhar por longas horas	-.55	-.15	.52	.17	1.98	2.01
25	Meu empregador me fornece as coisas que preciso para desenvolver meu trabalho	-1.12	-.42	.59	.94	.92	1.01
26	Meu ambiente de trabalho é seguro	-.77	-.31	.19	.89	.66	.65
27	As condições de trabalho são satisfatórias	-.40	-.43	.46	.37	.83	.82
28	Eu estou feliz com a infraestrutura do local em que trabalho	-.17	-.27	-.02	.46	.85	.80
29	A empresa que trabalho possui uma comunicação clara com os funcionários	-.31	.07	-.21	.46	2.17	3.06
30	Sinto orgulho em contar para as outras pessoas que faço parte da empresa que trabalho	-.72	-.13	-.01	.85	.75	.78
31	Eu recomendaria esta empresa como um bom local para se trabalhar	-.28	-.80	.53	.55	.79	.78

Note. Item numbering follows the scoring order of the WRQoL-2; (R) indicates reverse-scored items. Items with at least one disordered threshold: 5, 6, 9, 10, 18, 19, 21, 23, 24, 27, 28, 29, and 31. This study focuses on the evidence supporting the Brazilian Portuguese item wording; future studies are recommended to test the instrument in other languages. Licensing is under ©2024 QoWL (<https://www.qowl.co.uk/>).

The delta-tau parameterization of the PCM provided insights into the difficulty levels of each item's thresholds, which reflect the progression of the latent trait. Inspection of threshold ordering showed that 18 of the 31 items had monotonically increasing thresholds, as the model expects. In 13 items (5, 6, 9, 10, 18, 19, 21, 23, 24, 27, 28, and 31, in addition to item 29) at least one threshold was lower than the preceding one, which means that the corresponding intermediate category was never the most probable response at any level of the trait. Disordering was most pronounced in item 9 (−.39, .19, .16, .05), whose three upper thresholds were essentially collapsed, and in items 19 (−1.05, .37, .14, .54) and 29 (−.31, .07, −.21, .46). These are the same items that displayed the highest infit and outfit values, which indicates that their misfit is at least partly a matter of category functioning rather than of item content alone. Items with ordered thresholds spanning a wide range, such as item 1 (−.63, −.44, −.27, 1.34) and item 17 (−1.21, −.38, −.04, 1.60), behaved as the model expects.

Infit and outfit statistics were used to assess the fit of each item to the PCM. Most items demonstrated acceptable values, suggesting that they functioned as expected within the scale. However, five items displayed elevated values: item 7 (infit = 2.36, outfit = 2.73), item 9 (infit = 3.10, outfit = 4.20), item 19 (infit = 2.66, outfit = 3.15), item 24 (infit = 1.98, outfit = 2.01), and item 29 (infit = 2.17, outfit = 3.06). Conversely, items such as 2 (infit = .69, outfit = .67) and 22 (infit = .58, outfit = .57) displayed good fit statistics, indicating consistency in measuring the latent trait. Taken together, the item-level results indicate that the WRQoL-2 separates respondents reliably in this sample but that a substantial subset of items requires revision: five

items misfit the model and 13 items show malfunctioning response categories. Next, the likelihood ratio chi-square statistics for DIF are presented in Table 2.

Table 2

Likelihood Ratio Chi-Square Statistics for Differential Item Functioning (DIF)

Item	χ^2	<i>p</i>	<i>Adj. p</i>	Item	χ^2	<i>p</i>	<i>Adj. p</i>
1	2.655	.265	.601	17	0.241	.887	.921
2	0.837	.658	.921	18	3.409	.182	.559
3	3.236	.198	.559	19	4.802	.091	.401
4	2.760	.252	.601	20	0.285	.867	.921
5	0.837	.658	.921	21 (R)	2.513	.285	.601
6	0.959	.619	.921	22 (R)	0.374	.830	.921
7	5.733	.057	.316	23 (R)	0.295	.863	.921
8	5.591	.061	.316	24 (R)	1.124	.570	.921
9	12.240	.002	.036	25	6.895	.032	.316
10	12.109	.002	.036	26	0.327	.849	.921
11	4.297	.117	.452	27	2.472	.291	.601
12 (R)	0.936	.626	.921	28	0.813	.666	.921
13	0.254	.881	.921	29	0.117	.943	.943
14	1.105	.575	.921	30	3.761	.152	.525
15	5.643	.060	.316	31	0.229	.892	.921
16	0.673	.714	.921				

Note. (R) indicates reverse-scored items. *Adj. p* = *p*-value adjusted for multiple comparisons across the 31 items using the Holm step-down procedure (Holm, 1979).

Subgroups were defined as respondents without and with a diagnosis of mental health conditions. Most items did not show significant DIF; for example, item 2 ($\chi^2 = 0.84$, *p* = .658, *adj. p* = .921) and item 12 ($\chi^2 = 0.94$, *p* = .626, *adj. p* = .921). Item 9 ($\chi^2 = 12.24$, *p* = .002, *adj. p* = .036) and item 10 ($\chi^2 = 12.11$, *p* = .002, *adj. p* = .036) showed significant DIF after Holm adjustment. Item 25 was significant before adjustment ($\chi^2 = 6.90$, *p* = .032) but not after it (*adj. p* = .316), and items 7, 8, 15, and 19 had unadjusted *p*-values between .057 and .091. The predominance of non-significant results may indicate a tendency towards stable item functioning across subgroups, but this pattern is provisional: with only 41 respondents in the reference group, it requires confirmation in a more robust design before it can be read as evidence that the corresponding items are free of DIF (Belzak, 2020). Table 2 therefore identifies items 9 and 10 as requiring revision, alongside a provisional indication of stability in the remaining items.

Discussion

This study examined the item-level functioning of the WRQoL-2 in a Brazilian sample and produced a mixed picture. Person reliability was high (.930), indicating that the item set separates respondents along the latent continuum with acceptable precision (Lang et al., 2019). Four findings qualify this result: the assumption of local independence was not fully supported, 13 of the 31 items showed disordered thresholds, five items showed elevated infit and outfit values, and two items showed significant DIF. The instrument is therefore usable as a global measure in this population, but it is not yet a fully calibrated item bank, and the items flagged here require revision before the scale is used to compare workers with and without a mental health diagnosis.

The comparison between the PCM and the RSM did not yield a single preferred model. The likelihood ratio test favoured the PCM, but AIC, BIC, and CAIC all favoured the RSM, and the magnitude of these differences was substantial ($\Delta\text{BIC} = 332$). When a complexity-penalized criterion and an unpenalized test disagree in this way, and when the more complex model estimates 125 parameters from 209 respondents, the reasonable conclusion is that the data do not support item-specific threshold structures, rather than that the more complex model is superior (Raykov & Pusic, 2025). This reading is consistent with the estimation literature on small samples, which shows that item and threshold parameters in polytomous Rasch-family models become unstable and acquire large standard errors as the parameter-to-person ratio increases (Finch & French, 2019). The RSM is therefore treated here as the model preferred on statistical grounds, and the PCM as a diagnostic device. The practical value of the PCM in this study is that its item-specific thresholds make category malfunctioning visible: without them, the 13 items with disordered categories would not have been identified (Wind, 2023).

Item-Level Insights

Items 7 (“I feel able to voice opinions and influence changes in my area of work”) and 9 (“I am involved in decisions that directly affect members of the public”) demonstrated infit values of 2.36 and 3.10 and outfit values of 2.73 and 4.20. Values of this magnitude indicate that responses to these items were considerably more variable than the model predicts, which may reflect wording or context that is interpreted differently by different groups of respondents (Jones et al., 2023). For item 9, the combination of a very high outfit value with three collapsed upper thresholds suggests that the item is not operating as a five-category item in this sample: respondents at similar trait levels selected widely different categories, which is what would be expected if involvement in decisions affecting the public is simply not applicable to a substantial portion of a heterogeneous sample of workers. Items 19 (“My line manager actively promotes flexible hours/patterns”) and 29 (“The organisation communicates well with its employees”) also displayed elevated values (infit = 2.66 and 2.17; outfit = 3.15 and 3.06), which may reflect genuine variability in managerial practices and in communication across the organizations represented in the sample, particularly regarding the availability of flexible working arrangements (Branicki et al., 2024; Müller, 2020).

Differential Item Functioning (DIF) Analysis

Items 9 (“I am involved in decisions that directly affect members of the public”) and 10 (“I have sufficient opportunities to question managers about change at work”) showed significant DIF after Holm adjustment. Both refer to voice and participation in organizational decisions, which is the domain in which DIF was expected to concentrate. The remaining items did not reach significance, which may suggest a tendency towards invariance across subgroups; this reading, however, requires confirmation in a more robust design. With 41 respondents in the

group without a diagnosis, the likelihood ratio test has limited power for polytomous items, so the risk of Type II error is not negligible (Belzak, 2020). Table 2 is therefore best read as identifying two items with sufficient evidence to warrant revision, alongside a provisional pattern of stability in the remaining items that a larger and better balanced sample should confirm.

A further constraint concerns the direction of the DIF. The likelihood ratio test employed here is an omnibus test: it indicates that item parameters differ between groups but not whether the item is relatively more or less difficult for either group. Establishing direction would require estimating and comparing item locations and thresholds separately within each group, which was not performed in the present study and would in any case be poorly determined with 41 respondents in the smaller group. Consequently, no substantive interpretation is offered here as to how items 9 and 10 might disadvantage either subgroup; the finding reported is that their parameters are not invariant. Determining the direction and magnitude of this DIF, ideally with a purposefully recruited and better balanced sample, is a necessary next step before the items are rewritten (Russell et al., 2021).

Implications for Future Research and Practice

The location of the DIF in items concerning voice and participation is consistent with the literature on mental health stigma at work, which describes how employees with a diagnosis may perceive greater barriers to expressing opinions and influencing decisions (Østerud, 2023; Santuzzi et al., 2025). This remains a plausible interpretation rather than a demonstrated mechanism: the present design does not include measures of perceived stigma, psychological safety, or organizational support, so alternative explanations, such as differences in occupational role or in the applicability of the item content, cannot be ruled out. Testing the stigma interpretation would require designs that measure these constructs directly.

Cortez, P. A. et al., (2026). WRQoL-2: Differential Item Functioning Analysis

Items 7, 9, 10, 19, 24, and 29 may require revision to enhance their alignment with the measurement model. Potential revisions include clarifying item wording, adding a not-applicable option for items such as 9 whose content does not apply to all workers, and reconsidering the number of response categories for the 13 items with disordered thresholds (Santuzzi et al., 2022). Until such revisions are implemented and cross-validated, practitioners and researchers should not use items 9 and 10 to compare the work-related quality of life of workers with and without a mental health diagnosis.

Limitations

Five limitations should be considered when interpreting these results. First, the sample of 209 respondents is small for the calibration of a 31-item polytomous instrument and clearly small for the PCM, which estimates 125 parameters from these data, a parameter-to-person ratio of 0.60. Item and threshold estimates should therefore be regarded as provisional, and the model comparison reported here should not be read as a definitive statement about which parameterization best represents the WRQoL-2 (Finch & French, 2019). Second, the DIF analysis compared 168 respondents with a diagnosis to 41 without one. This imbalance limits statistical power, so the DIF results constitute a screening procedure rather than a test of measurement invariance, and the non-significant results obtained for 29 items are inconclusive (Belzak, 2020). The direction and magnitude of the DIF detected in items 9 and 10 also remain undetermined, since the omnibus likelihood ratio test does not provide this information.

Third, confidence intervals are not reported for person reliability or for the item and threshold parameters, so the precision of these estimates cannot be assessed from the results presented. Given the sample size, the standard errors of the PCM parameters are expected to be large. Future analyses of these data should report interval estimates, for example through

parametric bootstrapping. Fourth, dimensionality was not re-examined in the present analysis, and the unidimensionality on which the Rasch-family models rest is an approximation rather than a demonstrated property of these data. The confirmatory factor analysis previously conducted with this same sample (Cortez et al., 2025); the present calibration is therefore justified by the coherence of the total score - high common variance among items, high total-score reliability, and strong dimension-to-total correlations - and not by a formally estimated general factor. Consistent with this, the significant MADaQ3 obtained here indicates that residual dependence remains. Analyses that model the seven dimensions explicitly, such as multidimensional or bifactor IRT models, would represent the instrument more appropriately and are recommended for future work with larger samples.

Fifth, recruitment used snowball sampling through public organizational groups on social networks, and the resulting sample is not representative of Brazilian workers: 73.7% were women, 87.6% had completed a degree, 77.0% self-identified as white, and 80.4% belonged to the mental health group. The findings describe item functioning in this specific configuration and should not be generalized to the Brazilian working population. This also qualifies the framing of the study: the analyses contribute to the goal of socially fair assessment by identifying items that require revision, but a claim about fairness across diverse employee groups would require samples that actually represent that diversity, particularly with respect to education, race, and occupational sector (Santuzzi et al., 2022). Finally, the reliance on self-reported data and the cross-sectional design restrict causal inference, and detailed clinical records were not available.

Final Considerations

A question that remains open is whether the differential functioning observed in items concerning voice and participation is a property of the WRQoL-2 or reflects broader social and

Cortez, P. A. et al., (2026). WRQoL-2: Differential Item Functioning Analysis

organizational dynamics in Brazil, including stigma toward mental health and restricted opportunities for participation at work. The present design cannot separate these possibilities. What the study establishes is more specific: in this Brazilian sample, the WRQoL-2 separates respondents reliably, the parsimonious RSM is preferred over the PCM by all complexity-penalized criteria, 13 items show malfunctioning response categories, five items misfit the model, and two items concerning voice and participation function differently across groups defined by mental health status.

Two alternative readings of these item-level findings deserve consideration in future work, and they are not reducible to a problem of item content. The first is a model-level explanation: the misfit and the disordered thresholds may indicate that neither the PCM nor the RSM is the most appropriate representation of these data. Models that accommodate the seven correlated dimensions explicitly, or polytomous models that allow discrimination to vary across items, may represent the instrument better, and the preference for the parsimonious RSM observed here is consistent with the possibility that a simpler category structure is closer to the way Brazilian respondents actually use the scale.

The second is a measurement-context explanation: the five-category format and the wording of its anchors were developed in a different linguistic and organizational setting, so the disordering may reflect categories that are insufficiently distinguishable for Brazilian respondents rather than a defect in item content. If this is the case, reducing or relabelling categories and calibrating the anchors to the conventions of Brazilian work settings would be a more economical route to improvement than rewriting the items. Both readings are testable, and distinguishing between them is a productive next step. Accordingly, the items flagged here should be revised, and the instrument should be re-examined in larger, better balanced, and more representative

Paidéia, 36, e3628

samples before it is used for subgroup comparisons. Within these boundaries, the study contributes the first item-level map of the Brazilian WRQoL-2 and identifies precisely where refinement is needed.

References

- Belzak, W. C. M. (2020). Testing differential item functioning in small samples. *Multivariate Behavioral Research*, 55(5), 722-747. <https://doi.org/10.1080/00273171.2019.1671162>
- Branicki, L. J., Brammer, S., Brosnan, M., Lazaro, A. G., Lattanzio, S., & Newnes, L. (2024). Factors shaping the employment outcomes of neurodivergent and neurotypical people: Exploring the role of flexible and homeworking practices. *Human Resource Management*, 63(6), 1001-1023. <https://doi.org/10.1002/hrm.22243>
- Christensen, K. B., Makransky, G., & Horton, M. (2017). Critical values for Yen's Q3: Identification of local dependence in the Rasch model using residual correlations. *Applied Psychological Measurement*, 41(3), 178-194. <https://doi.org/10.1177/0146621616677520>
- Cortez, P. A., Veiga, H. M. S., Moscon, D. C. B., Mendonça, H., Neiva, E. R., & Van Laar, D. (2025). Evidências de validade do instrumento qualidade de vida no trabalho 2 (WRQoL-2) no contexto brasileiro [Validity evidence of the Work-Related Quality of Life 2 (WRQoL-2) instrument in the Brazilian context]. *Psico*, 56(1), e47152. <https://doi.org/10.15448/1980-8623.2026.1.47152>
- Easton, S., & Van Laar, D. (2013). QoWL (Quality of Working Life)-What, how, and why? *Journal of Psychology Research*, 3(10), 596-605. <https://doi.org/10.17265/2159-5542/2013.10.006>

Cortez, P. A. et al., (2026). WRQoL-2: Differential Item Functioning Analysis

- Easton, S., & Van Laar, D. (2018). *User manual for the Work-Related Quality of Life (WRQoL) scale: A measure of quality of working life*. University of Portsmouth.
- Finch, H., & French, B. F. (2019). A comparison of estimation techniques for IRT models with small samples. *Applied Measurement in Education*, 32(2), 77-96.
<https://doi.org/10.1080/08957347.2019.1577243>
- Garzaro, G., Clari, M., Donato, F., Dimonte, V., Mucci, N., Easton, S., Van Laar, D., Gatti, P., & Pira, E. (2020). A contribution to the validation of the Italian version of the Work-Related Quality of Life Scale. *La Medicina del Lavoro*, 111(1), 32-45.
<https://doi.org/10.23749/mdl.v111i1.8570>
- Hagquist, C., & Andrich, D. (2017). Recent advances in analysis of differential item functioning in health research using the Rasch model. *Health and Quality of Life Outcomes*, 15, 181.
<https://doi.org/10.1186/s12955-017-0755-0>
- Hess, J. (2024). Bipolar in the academy: A case of testimonial smothering. *Disability Studies Quarterly*, 43(2). <https://doi.org/10.18061/dsq.v43i2.8491>
- Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6(2), 65-70. <https://www.jstor.org/stable/4615733>
- Hrnjicic, A., & Alihodžic, A. (2024). Measuring students' conceptual understanding of real functions: A Rasch model analysis. *International Electronic Journal of Mathematics Education*, 19(1), em0761. <https://doi.org/10.29333/iejme/13942>
- Jones, E. A., Wind, S. A., Tsai, C.-L., & Ge, Y. (2023). Comparing person-fit and traditional indices across careless response patterns in surveys. *Applied Psychological Measurement*, 47(5-6), 365-385. <https://doi.org/10.1177/01466216231194358>

Paidéia, 36, e3628

- Lang, J. W. B., Lievens, F., De Fruyt, F., Zettler, I., & Tackett, J. L. (2019). Assessing meaningful within-person variability in Likert-scale rated personality descriptions: An IRT tree approach. *Psychological Assessment*, 31(4), 474-487. <https://doi.org/10.1037/pas0000600>
- Li, J.-N., Chen, X.-Q., Li, Q.-Q., Zheng, Q.-X., Pan, Y. Q., Huang, L., Zhu, Y., Liu, R.-L., & Jiang, X.-M. (2023). Effect of depressive symptoms on quality of work life in female nurses: A cross-sectional study using propensity score matching. *Frontiers in Psychiatry*, 14, 1213025. <https://doi.org/10.3389/fpsyt.2023.1213025>
- Li, P., Wang, Y., & Zhang, M. (2022). Translation and validation of the Work-Related Quality of Life Scale (WRQoLS-2) in a nursing cohort. *Contemporary Nurse*, 58(5-6), 435-445. <https://doi.org/10.1080/10376178.2022.2147849>
- Müller, M. (2020). Item fit statistics for Rasch analysis: Can we trust them? *Journal of Statistical Distributions and Applications*, 7, 5. <https://doi.org/10.1186/s40488-020-00108-7>
- Olteanu, A., Castillo, C., Diaz, F., & Kıcıman, E. (2019). Social data: Biases, methodological pitfalls, and ethical boundaries. *Frontiers in Big Data*, 2, 13. <https://doi.org/10.3389/fdata.2019.00013>
- Østerud, K. L. (2023). Mental illness stigma and employer evaluation in hiring: Stereotypes, discrimination and the role of experience. *Sociology of Health & Illness*, 45(1), 90-108. <https://doi.org/10.1111/1467-9566.13544>
- Peres, R. S., & Cortez, P. A. (2024). Psychosocial impact of the COVID-19 pandemic in Brazilian post-peak period: Differences between individuals with and without pre-existing psychiatric conditions. *International Journal of Environmental Research and Public Health*, 22(1), 27. <https://doi.org/10.3390/ijerph22010027>

Cortez, P. A. et al., (2026). WRQoL-2: Differential Item Functioning Analysis

- Raykov, T., & Pusic, M. (2025). A note on evaluation of polytomous item locations with the rating scale model and testing its fit. *Educational and Psychological Measurement*, 85(3), 527-538. <https://doi.org/10.1177/00131644241259026>
- Russell, M., Szendey, O., & Kaplan, L. (2021). An intersectional approach to DIF: Do initial findings hold across tests? *Educational Assessment*, 26(4), 284-298. <https://doi.org/10.1080/10627197.2021.1965473>
- Sabino, A., Moreira, A., Cesário, F., & Pinto-Coelho, M. (2024). Adaptation of the Work-Related Quality of Life-2 Scale (WRQoL-2) among Portuguese workers. *Societies*, 14(7), 120. <https://doi.org/10.3390/soc14070120>
- Santuzzi, A. M., Martinez, J. J., & Keating, R. T. (2022). The benefits of inclusion for disability measurement in the workplace. *Equality, Diversity and Inclusion: An International Journal*, 41(3), 474-490. <https://doi.org/10.1108/EDI-06-2020-0167>
- Santuzzi, A. M., Keating, R. T., & Martinez, J. J. (2025). The impact of response options on formal disclosure rates for disabilities in organizations. *Group & Organization Management*, 50(4), 1097-1110. <https://doi.org/10.1177/10596011241258669>
- Sezer, B., Sıddikoğlu, D., Doğan, İ., Taşdemir, M. C., & Ateş, G. (2024). The influence of sociodemographic and health-related factors, clinical education status, and sleep quality on work-related quality of life among dental students. *European Journal of Dental Education*, 28(3), 806-815. <https://doi.org/10.1111/eje.13010>
- Silarova, B., Brookes, N., Palmer, S., Towers, A.-M., & Hussein, S. (2022). Understanding and measuring the work-related quality of life among those working in adult social care: A scoping review. *Health & Social Care in the Community*, 30(5), 1637-1664. <https://doi.org/10.1111/hsc.13718>

Paidéia, 36, e3628

Sinval, J., Sirgy, M. J., Lee, D.-J., & Marôco, J. (2020). The quality of Work Life Scale: Validity evidence from Brazil and Portugal. *Applied Research in Quality of Life*, 15, 1323-1351.

<https://doi.org/10.1007/s11482-019-09730-3>

Van Vo, B., Sunantiwat, M., Puntong, S., Vo, T. Q., & Anuratpanich, L. (2024). A cross-sectional study of the quality of work life of hospital pharmacists in southern Vietnam. *Journal of Pharmaceutical Health Services Research*, 15(3), rmae015.

<https://doi.org/10.1093/jphsr/rmae015>

Wang, Y., Byrne, L., Bartram, T., & Chapman, M. (2023). Developing inclusive and healthy organizations by employing designated lived experience roles: Learning from human resource management innovations in the mental health sector. *The International Journal of Human Resource Management*, 34(10), 1973-2001.

<https://doi.org/10.1080/09585192.2022.2054287>

Wind, S. A. (2023). Detecting rating scale malfunctioning with the partial credit model and generalized partial credit model. *Educational and Psychological Measurement*, 83(5), 953-983. <https://doi.org/10.1177/00131644221116292>

Pedro Afonso Cortez is a Professor of the Universidade Federal de Uberlândia, Uberlândia-MG, Brazil.

Heila Magali da Silva Veiga is a Professor of the Universidade Federal de Uberlândia, Uberlândia-MG, Brazil.

Helenides Mendonça is a Professor of the Pontifícia Universidade Católica de Goiás, Goiânia-GO, Brazil.

Elaine Rabelo Neiva is a Professor of the Universidade de Brasília, Brasília-DF, Brazil.

Darren Van Laar is a Professor of the University of Portsmouth, Portsmouth, United Kingdom.

Cortez, P. A. et al., (2026). WRQoL-2: Differential Item Functioning Analysis

Research Data Availability

The datasets generated and/or analyzed during the current study are available from the corresponding author on reasonable request.

Conflict of interest

The authors have no conflicts of interest to declare.

Declaration of Use of Artificial Intelligence

The authors declare that no artificial intelligence tools were used in the writing or editing of this manuscript.

Authors' Contribution:

All authors made substantial contributions to the conception and design of this study, to data analysis and interpretation, and to the manuscript revision and approval of the final version. All the authors assume public responsibility for the content of the manuscript.

Associate editor:

Cristiane Faiad Moura

Received: Aug. 18, 2025

1st Revision: Aug. 13, 2026

Approved: Aug. 20, 2026

How to cite this article:

Paidéia, 36, e3628

Cortez, P. A., Veiga, H. M. S., Mendonça, H., Neiva, E. R., & Van Laar, D. (2026). Work-Related Quality of Life Scale (WRQoL-2): Differential Item Functioning Analysis. *Paidéia (Ribeirão Preto)*, 36, e3628. <https://doi.org/10.1590/1982-4327e3628>

This preprint was submitted under the following conditions:

- The authors declare that the necessary Terms of Free and Informed Consent of participants or patients in the research were obtained and are described in the manuscript, when applicable.
- The authors declare that the preparation of the manuscript followed the ethical norms of scientific communication.
- The authors declare that they are aware that they are solely responsible for the content of the preprint and that the deposit in SciELO Preprints does not mean any commitment on the part of SciELO, except its preservation and dissemination.
- The authors declare that the data, applications, and other content underlying the manuscript are referenced.
- The deposited manuscript is in PDF format.
- The authors declare that the research that originated the manuscript followed good ethical practices and that the necessary approvals from research ethics committees, when applicable, are described in the manuscript.
- The authors declare that once a manuscript is posted on the SciELO Preprints server, it can only be taken down on request to the SciELO Preprints server Editorial Secretariat, who will post a retraction notice in its place.
- The authors agree that the approved manuscript will be made available under a [Creative Commons CC-BY](#) license.
- The submitting author declares that the contributions of all authors and conflict of interest statement are included explicitly and in specific sections of the manuscript.
- The authors declare that the manuscript was not deposited and/or previously made available on another preprint server or published by a journal.
- If the manuscript is being reviewed or being prepared for publishing but not yet published by a journal, the authors declare that they have received authorization from the journal to make this deposit.
- The submitting author declares that all authors of the manuscript agree with the submission to SciELO Preprints.