

Article

Not peer-reviewed version

Discrepancies in the Country Versions of the WHOQOL-BREF as a Potential Source of Error in Assessing Quality of Life and a Barrier to Comparative Research

[Stanisław Maksymowicz](#)^{*}, [Michał B. Paradowski](#), Maria Libura, [Andrzej Jarynowski](#)

Posted Date: 15 December 2025

doi: 10.20944/preprints202512.1223.v1

Keywords: health-related quality of life; WHOQOL-BREF questionnaire; translation; comparability; methodology



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Discrepancies in the Country Versions of the WHOQOL-BREF as a Potential Source of Error in Assessing Quality of Life and a Barrier to Comparative Research

Stanisław Maksymowicz ^{1,*}, Maria Libura ², Andrzej Jarynowski ³ and Michał B. Paradowski ⁴

¹ Department of Psychology and Sociology of Health and Public Health, School of Public Health, Collegium Medicum, University of Warmia and Mazury in Olsztyn, Poland

² Chair of Epidemiology and Preventive Medicine, Faculty of Medicine, Jagiellonian University Medical College in Krakow, Poland

³ System Modeling Group, Institute of Veterinary Epidemiology and Biostatistics, Freie Universität Berlin, 14163 Berlin, Germany

⁴ Institute of Applied Linguistics, University of Warsaw, Poland

* Correspondence: smaksymowicz@gmail.com

Abstract

Purpose: The aim of the study is to determine whether different language versions of the WHOQOL-BREF questionnaire allow for obtaining comparable health-related quality of life measurement results across different countries. **Methods:** We analyze 17 country versions published on the World Health Organization website. Each version was compared to the reference English questionnaire in terms of translation, timeframe of measurement, inclusion of instructions for respondents, visual layout and additional elements. To evaluate the accuracy of the questions and scales, selected versions were reviewed by native speakers and all translations were machine-translated and subsequently reviewed using a large language model tool. **Results:** The results reveal substantial discrepancies between the versions across all evaluated criteria. In some cases, translation errors were so severe that responses to certain questions could not be meaningfully compared. **Conclusions:** We conclude that the WHOQOL-BREF questionnaire translations should undergo a comprehensive review to ensure conceptual and structural consistency and comparability. Additionally, language versions developed by external institutions should be carefully vetted by the WHO. The common practice of adding a note that a certain version is not an official WHO translation is insufficient, as these tools are frequently treated by researchers and clinicians as *de facto* official instruments.

Keywords: health-related quality of life; WHOQOL-BREF questionnaire; translation; comparability; methodology

Introduction

Health-related quality of life (HRQoL) research is a critical component of many clinical studies and medical experiments. It is an integral part of numerous analyses, helping understand the impact of medical interventions not only in their physiological aspects, but also in mental and social domains, which are integral to evaluating therapeutic success.

HRQoL can be assessed in various ways. In its simplest form, it involves asking the patient in an open interview about their subjective assessment of their situation. However, to obtain data that is as objective as possible and suitable for statistical analysis and comparison, standardized tools must be used. This role is fulfilled by health-related quality of life assessment questionnaires.

In clinical practice, researchers typically use tools with simplified structures to obtain easily analyzable data. One such tool is the WHOQOL-BREF [1], a shortened version of the health-related quality of life assessment questionnaire developed by the World Health Organization. The WHOQOL-BREF consists of 26 questions that address four domains of everyday life functioning: physical, psychological, social, and environmental. It is a shortened version of the WHOQOL-100 [2], making it easier to use. Each question employs a five-point Likert scale. The questionnaire is used in both clinical practice and scientific research. As of July 11, 2025, there were 3,812 mentions on WHOQOL-BREF [Title/Abstract] in PubMed and over 74,600 in Google Scholar.

The WHOQOL-BREF is a practical tool translated into many languages. However, while using the Polish version in our research, we identified numerous structural and translation errors. This led us to investigate whether similar issues exist in other language versions, the gravity of the discrepancies, and whether they could compromise the reliability and comparability of responses across countries. The aim of this study is therefore to assess the extent of these discrepancies across selected translations and evaluate whether the tool's international comparability.

Materials and Methods

For this analysis, we reviewed 17 language versions from the WHO website of the WHOQOL-BREF questionnaires in PDF and WORD formats. The languages included Australian English, Bulgarian, Czech, Danish, Dutch (Netherlands), Finnish, French, German, Greek, Italian, Lithuanian, Norwegian, Polish (including a revised Polish version we had improved), Romanian, Slovak, Spanish, and Swedish. All represent European countries, with Australian version added to enable comparison between Australian English and reference English version from the official WHO manual (https://www.who.int/tools/whoqol/whoqol-bref/docs/default-source/publishing-policies/whoqol-bref/english_whoqol_bref).

We conducted a qualitative analysis of translation quality for both individual questions and response scales, using two approaches: (1) review by native speakers (for selected questionnaires in Danish, Dutch, Finnish, German, Polish, and Swedish), and (2) assessment using a large language model (LLM) for all versions (ChatGPT 4o, comparing differences between the official English version and the 17 language versions, translated with machine translation tool into English). Furthermore, all questionnaires were analyzed manually in terms of timeframe respondents were instructed to consider when answering, inclusion of essential elements such as instructions for respondents, visual layout and additional components (see Figure 1 and Table 1).

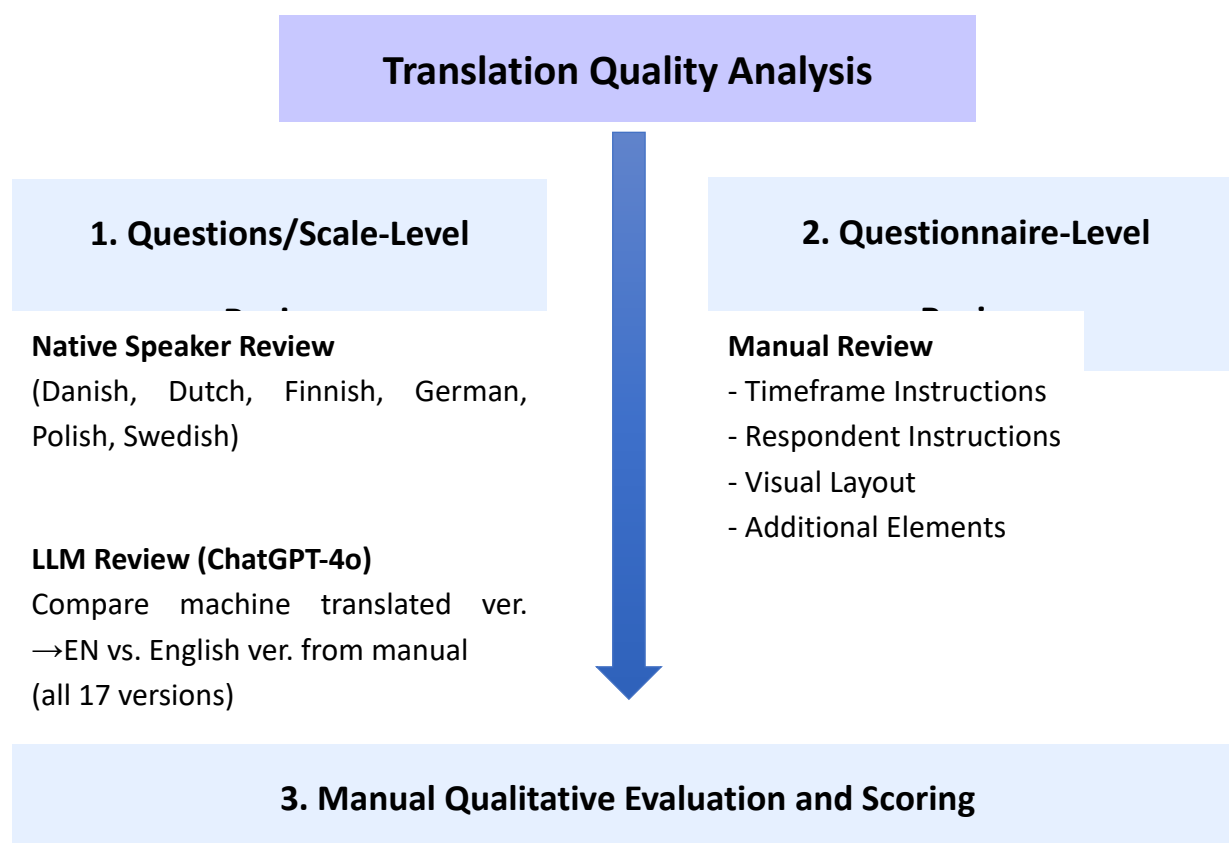


Figure 1. Translation quality methodology.

All versions were compared to the reference English questionnaire provided as an annex to the official WHO manual. The software used for the analyses included ChatGPT 4o and Google Translate.

The degree of deviation from the original translation was assessed on 9 levels, each assigned a corresponding score with weighted factors (for instance, an incorrect measurement timeframe received higher weight than the missing demographic fields). This permitted the calculation and comparison of the magnitude of deviations across versions (see Table 1). For questions and response scales, we only assessed the general presence of linguistic discrepancies, not deviations due to country-specific adaptation (such as the minor differences in scale questions observed between the reference English version and Australian English version). A higher score indicates a greater divergence from the reference/original.

Table 1. Criteria for assessing individual areas.

| No | Evaluated Area | Deviation - Scoring | Explanation | Analysis Method* |
|----|---------------------|-------------------------------------|---|------------------|
| 1. | Visual Presentation | 0 (no deviation) 1 (deviation) | Different visual layout compared to the original | R |
| 2. | Demographics | 0 (no deviation) 0.5 (deviation) | Missing demographic section at the beginning of the questionnaire | R |
| 3. | Title | 0 (no deviation) 0.5 (deviation) | Title missing or altered | R |

| | | | | |
|----|---|--|--|-------|
| 4. | Introduction and Instructions for the Patient | 0 (no deviation) 1 (deviation) | Missing instructions at the beginning | R |
| 5. | Timeframe for Measurement | 0 (no deviation) 2 (deviation) | Timeframe differs from the intended two-week period | R |
| 6. | Questions | 0 (no significant deviation) 1 (deviation in 1 question) 2 (deviation in more than 1 question) | If a deviation affecting item meaning is found in one question, 1 point is assigned; if multiple questions are affected, 2 points are assigned | R+C+N |
| 7. | Response scales | 0 (no significant deviation) 1 (deviation in 1 scale) 2 (deviation in more than 1 scale) | If a deviation is found in one scale that may alter its meaning, 1 point is assigned; if multiple scales are affected, 2 points are assigned | R+C+N |
| 8. | Acknowledgment and Summary | 0 (no deviation) 0.5 (deviation) | Missing formal acknowledgment at the end of the form and/or summary questions | R |
| 9. | Additional Elements Not Present in the Original | 0 (no deviation) 0.5 (deviation) | Extra content (e.g., scoring table) included at the beginning or end of the form | R |

*Analysis method legend: R: researcher, C: LLM and machine translation tool, N: native-speaker.

Results

1. Visual presentation and measurement timeframe

The first analyzed aspect concerned the visual presentation of the entire questionnaire and its questions: general layout, the existence of a demographic section, appropriate title, introduction and instructions for the respondent, acknowledgments, and additional elements. This area exhibited the most discrepancies, although it is important to note that, aside from the instructions for the respondent (the absence of instructions in 5 of the analyzed language versions could lead to errors from the respondent), these elements generally do not significantly impact the results. However, there are exceptions, such as when a question and its scale are split across two pages, which can greatly hinder completing the questionnaire (this was the case in the Slovak and French versions, where a portion of a question in the PDF file—intended for printing by medical staff—was divided, see Figure 2).

There were also issues with the readability of fonts and colors and graphic design, as in the Italian version, where an illegible italic font and heavy borders reduced readability and visual clarity (Figure 4).

| |
|--|
| <p>2 (G4) E' soddisfatto/a della Sua salute ? <i>1 Molto insoddisfatto/a 2 Insoddisfatto/a 3 Nè soddisfatto/a nè insoddisfatto/a 4 Soddisfatto/a 5 Molto soddisfatto/a</i></p> |
| <p>Nelle domande seguenti, Le viene chiesto <u>in che misura</u> Lei ha sperimentato determinate cose negli ULTIMI 15 GIORNI.</p> |
| <p>3 (F1.4) In che misura i dolori fisici Le impediscono di fare le cose che deve fare ? <i>1 Per niente 2 Poco 3 Abbastanza 4 Molto 5 Moltissimo</i></p> |
| <p>4 (F11.3) Ha bisogno di trattamenti o interventi medici per poter affrontare la vita di tutti i giorni ? <i>1 Per niente 2 Poco 3 Abbastanza 4 Molto 5 Moltissimo</i></p> |
| <p>5 (F4.1) Quanto si gode la vita? <i>1 Per niente 2 Poco 3 Abbastanza 4 Molto 5 Moltissimo</i></p> |
| <p>6 (F24.2) In che misura Lei pensa che la Sua vita abbia un significato? <i>1 Per niente 2 Poco 3 Abbastanza 4 Molto 5 Moltissimo</i></p> |

Figure 4. Problematic formatting in the Italian version of the questionnaire (<https://www.who.int/tools/whoqol/whoqol-bref/docs/default-source/publishing-policies/whoqol-bref/italian-whoqol-bref>, author: Centro Collaborativo Italiano).

The most critical error, however, was related to the measurement intervals over which the patient was asked to assess their experiences. The original version of the questionnaire instructs the patient to reflect on the past two weeks, a period that is easy to conceptualize. However, six of the analyzed language versions used different periods: four weeks in the Czech and Polish (unrevised) versions, a month in the Lithuanian and Norwegian versions, and 15 days in the Italian version. The French version did not specify a time interval at all. This means that in these cases, respondents were asked to answer based on a different timeframe than originally intended, which could impact the results, undermining the comparability of data, especially in cross-national analyses that assume methodological uniformity.

2. Translation errors in questions and response scales

The second and more critical area of analysis involved discrepancies in the translation of individual questions and response scales. While some differences may reflect appropriate cultural adaptation, many were clearly translation errors that compromise construct validity.

We began with a detailed review of the Polish version (before revision). One major concern was the consistent exclusive use of the masculine formal address (“Pan”, equivalent to “Mr.”), failing to account for non-male respondents. In Polish, respondents and patients are addressed as “Pan” or “Pani” (Mr. or Ms., not by first name, in line with Polish cultural norms). If the tool is intended for both genders, a direct form such as second-person singular verbs or a dual-form address such as “Pan/i” (“Sir/Madam”) should be used.

More seriously, certain questions in the first Polish version did not correspond to the originally intended domain. For instance, question 15, which in the English original reads “How well are you able to get around?” and refers to Domain 1 (physical – mobility), was translated as “Jak odnajduje się Pan w tej sytuacji?” meaning “How do you find yourself in this situation?”, which is a completely different domain. This difference cannot be explained by cultural adaptation during translation. In this case, the question addresses an entirely different, vague psychological domain, but in the end it will be calculated as mobility, severely skewing the analysis of the patient’s quality of life measurement. A similar mistranslation occurred in the Slovak version, where the question was rendered as: “Ako dobre vyjdete?” which translates to “How well do you get along?”. Again, this does not capture the concept of physical mobility, making the item invalid for its intended purpose.

Another illustrative example of clearly incorrect translation comes from the French questionnaire. Where in the English question 6 original reads “To what extent do you feel your life to be meaningful?” addressing the subjective sense of life’s meaning and purpose,, the French version shifts the focus to philosophical, religious or spiritual belief systems: “Vos croyances personnelles donnent-elles un sens à votre vie?” meaning “Do your personal beliefs give meaning to your life?” Such a discrepancy introduces a different conceptual framework and is difficult to justify based on cultural differences considered during validation. Another example is a slight change of concepts in the Dutch version, where the original question referring to the physical environment has been translated as “de gezondheid van uw omgeving” (“the health of your environment”). This rewording implies ecological or environmental health rather than accessibility, safety, or infrastructure, again leading to a divergence from the intended construct.

Unfortunately, similar issues were identified in most language versions. In some cases, the deviation may appear subtle but could nonetheless introduce systematic bias, particularly in large comparative studies.

The qualitative analysis also revealed problems in the translation of the response scales. Among others, there were grammatical errors in one of the scales in Polish, unnecessary modifications in the Australian version not justified by linguistic norms or validation evidence, and incorrect translations of scale anchors in the Danish and Italian versions.

These issues collectively undermine the conceptual and semantic equivalence required for meaningful cross-cultural comparison. A comprehensive breakdown with cumulative deviation scores and detailed analysis across evaluation criteria between the translated questionnaires are presented in Figure 5 and Table 2. The extent of discrepancies was highest in the Polish unrevised version (9.5, dropping to 0.5 after the revision) and lowest in Bulgarian and Spanish versions.

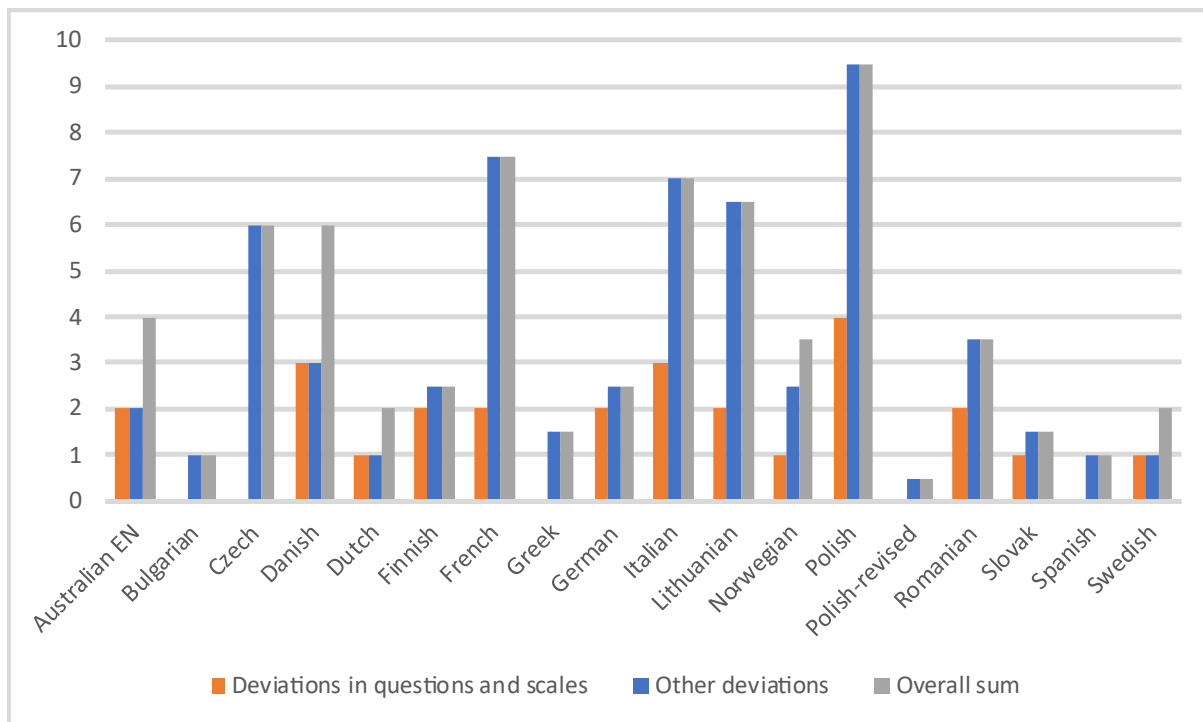


Figure 5. Deviations in individual language versions.

Table 2. Detailed score analysis.

| | Visual presentation | Demographics | Title | Introduction and instructions for the patient | Measurement time interval | Acknowledgements and Summary | Additional elements other than in the original | Questions | Scales | Sum of deviations without questions and scales | Overall sum |
|---------------|----------------------------------|------------------------------------|------------------------------------|---|----------------------------------|------------------------------------|--|--|--|--|-------------|
| | 0 (no deviation) - 1 (deviation) | 0 (no deviation) - 0.5 (deviation) | 0 (no deviation) - 0.5 (deviation) | 0 (no deviation) - 1 (deviation) | 0 (no deviation) - 2 (deviation) | 0 (no deviation) - 0.5 (deviation) | 0 (no deviation) - 0.5 (deviation) | 0 (no significant deviation) - 1 (deviation in 1 question). 2 (deviation in 2 or more questions) | 0 (no significant deviation) - 1 (deviation in 1 scale). 2 (deviation in 2 or more scales) | | |
| Australian EN | 0 | 0.5 | 0 | 0 | 0 | 1 | 0.5 | 0 | 2 | 2 | 4 |
| Bulgarian | 0 | 0.5 | 0.5 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 |
| Czech | 1 | 0.5 | 0.5 | 1 | 2 | 0.5 | 0.5 | 0 | 0 | 6 | 6 |
| Danish | 1 | 0.5 | 0 | 1 | 0 | 0.5 | 0 | 1 | 2 | 3 | 6 |
| Dutch | 0 | 0 | 0 | 0 | 0 | 0.5 | 0.5 | 1 | 0 | 1 | 2 |

| | | | | | | | | | | | |
|---------------------|---|-----|-----|---|---|-----|-----|---|---|-----|-----|
| Finnish | 0 | 0 | 0 | 0 | 0 | 0 | 0.5 | 2 | 0 | 2.5 | 2.5 |
| French | 1 | 0.5 | 0 | 1 | 2 | 0.5 | 0.5 | 2 | 0 | 7.5 | 7.5 |
| Greek | 1 | 0 | 0 | 0 | 0 | 0 | 0.5 | 0 | 0 | 1.5 | 1.5 |
| German | 0 | 0 | 0 | 0 | 0 | 0.5 | 0 | 2 | 0 | 2.5 | 2.5 |
| Italian | 1 | 0.5 | 0 | 0 | 2 | 0 | 0.5 | 2 | 1 | 7 | 7 |
| Lithuania n | 0 | 0.5 | 0 | 1 | 2 | 0.5 | 0.5 | 2 | 0 | 6.5 | 6.5 |
| Norwegia n | 0 | 0.5 | 0 | 0 | 2 | 0 | 0 | 1 | 0 | 2.5 | 3.5 |
| Polish | 1 | 0.5 | 0.5 | 1 | 2 | 0 | 0.5 | 2 | 2 | 9.5 | 9.5 |
| Polish – revised | 0 | 0.5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.5 | 0.5 |
| Romanian | 1 | 0 | 0 | 0 | 0 | 0 | 0.5 | 2 | 0 | 3.5 | 3.5 |
| Slovak | 0 | 0 | 0 | 0 | 0 | 0 | 0.5 | 1 | 0 | 1.5 | 1.5 |
| Spanish | 0 | 0.5 | 0.5 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 |
| Swedish | 0 | 0.5 | 0 | 0 | 0 | 0 | 0.5 | 1 | 0 | 1 | 2 |

Discussion

The advancement and globalization of science and medical research have increased the need for culturally adapted quality of life measurement tools. Where available, the adaptation of existing instruments is usually preferred over creating new ones, as this avoids the creation of superfluous tools [3], permits comparability of the results across cultures, and promotes information exchange in the scientific community [4].

However, translating subjective, patient-reported data assessment instruments involves more than simply rendering the text into another language [5–7]. The goal of translation is ensuring that the new version will maintain content integrity and conceptual consistency by measuring the target construct as closely as the original [8–11]. This expectation in cross-cultural research is commonly referred to as *construct equivalence* [12–14], the assumption that the tools employed will measure the same constructs in the same way whether in the original language/culture or in the translation [15–22].

Despite this, numerous researchers have highlighted a shortage of universally accepted, standardized methodological guidelines for questionnaire translation and cross-cultural adaptation [19,23–26]. While literature often details the translation process – that of the transformation of the questionnaire, many studies neglect the validation—or quality assessment—phase, which is critical for assessing translation quality [27]. Both European [15,28] and US regulatory bodies [29] have likewise expressed concerns over the methodological rigor and validity of translated tools. A methods review of 47 articles in cross-cultural nursing research by Maneesriwongul and Dixon (2004) found that the translation process was often inadequate [30]. Similarly, Hawkins et al. (2020) identified ten common types of translation errors across different versions of the Health Literacy Questionnaire [31]. Such errors may significantly compromise data interpretation. Our findings align with these prior critiques of translation methodologies.

Therefore, to ensure that questionnaire adaptations are reliable, valid, conceptually and semantically equivalent, and culturally relevant for target populations, several experts have advocated for rigorous robust multi-stage processes, from preliminary translation and cross-cultural adaptation [12,31–34] to subsequent validation [3,10,16,25,30,35–43]. Proper validation is vital for ensuring that the psychometric properties of the new version are not inferior to the original, helps guarantee that the information collected about the target population will be accurate, and enhances

acceptance of the questionnaire by professionals and society, helping reassure respondents' subjective perceptions, especially in the case of complex or sensitive questions [16,44,45].

To date, numerous translations of various iterations of the WHOQOL-BREF have been produced, spanning languages ranging from Akan [46] and Amharic [47] to Urdu [48] and Yoruba [49]. While most maintain their reliability and validity, some report low internal consistency in at least some of the measures [50–53].

This suggests that while the WHOQOL-BREF demonstrates broad cross-cultural utility, certain linguistic and cultural adaptations may negatively impact its psychometric properties, highlighting the need for careful contextual validation in each target population. Our study identifies some of the most critical vulnerabilities in the translation and implementation process and casts light on the scale of the problem, pointing to a need for greater oversight, harmonization, and quality assurance. This includes developing a robust translation protocol, providing training for national teams, and validating tools in collaboration with local and international experts, ideally in coordination with the WHO. This will help better ensure methodological consistency and facilitate meaningful cross-national comparisons of quality of life outcomes.

Conclusions

Our analysis reveals substantial discrepancies among the national versions of the WHOQOL-BREF tool, undermining the comparability of cross-country results. To address this issue, the instrument must undergo rigorous multilingual validation and standardization. The WHO should assume active oversight of the translation process to help safeguard linguistic accuracy and structural consistency. Furthermore, the organization should enforce a uniform layout and formatting standard across all language versions.

Although WHO mentions on its website and in the questionnaires that the translations were not conducted by the organization itself, the mere presence of these tools in the official WHO repository gives them the appearance of being formally endorsed. Researchers and clinicians frequently print and hand them out to patients and publish in medical journals without modification.

Unfortunately, our analysis shows that elements such as layout, instructions, additional elements, question wording, and response scale labels often vary significantly between versions. These translation differences may stem from two factors: (1) appropriate cultural adaptation and localization, and (2) translation errors. In the case of proper adaptation, supported by appropriate multi-step procedures (methodological rigor, piloting, psychometric validation,), variation may be acceptable and documented. However, translation errors can compromise the accuracy and validity of the results.

This issue requires urgent attention, as numerous studies based on the current versions of the WHOQOL-BREF may lack cross-national comparability, potentially distorting conclusions in research over many years. For example, a study might conclude that Polish patients have higher mobility than those in other countries, while in reality the item used measured how they assess their current situation. We recommend that WHO thoroughly review all existing translations, including both linguistic and structural elements, with particular attention to the preservation of instructions for the respondent and conceptual integrity. The organization's current disavowal of responsibility for this important tool negatively affects the quality of global research and its potential use in international contexts.

Additionally, the first author wishes to note that based on the findings of this study, a revised version of the Polish scale has been accepted and published by the WHO on their website. However, this revision addresses only one language version. The broader issue of the lack of comparability of the scales in other languages remains unresolved. We hope that this publication brings greater awareness to the problem and contributes to the advancement of more reliable comparable research in the field of quality of life assessment.

Limitations

The linguistic analysis of the selected questionnaires was conducted by native speakers, whose task was to verify the correctness of the translation. It must be noted that, due to the absence of external funding, these contributors were volunteers, not professional translators. As such, they mostly identified unnatural or awkward phrasing, which was not necessarily direct errors that would affect construct validity or the ability to accurately assess health-related quality of life.

To complement this subjective evaluation, we employed machine translation and a large language model to translate each version back into English and compare it with the original. This approach allowed us to identify several significant semantic discrepancies. Nevertheless, it is still possible that there are more errors that could only be detected by linguistically trained or psychometric experts with specialized training in questionnaire design and validation.

In the present study, we primarily focused on the problems in the Polish version—the questionnaire we eventually revised—and presented the most noticeable translation errors identified during machine translation in other language versions. Although this may appear limited in scope, it still has epistemic value as it highlights the broader issue of inconsistencies between different language versions, which was the purpose of this study.

Finally, only 17 language versions were included in the comparison. While this may seem restrictive, it should be emphasized that the goal of this work was not an exhaustive audit of all the available translations, which should be the task of the WHO, but to highlight the existing problem of using insufficiently standardized tools for international comparisons. The analysis of these selected tools alone revealed challenges that need to be addressed. In our opinion, a comprehensive review should be initiated by the WHO, and new versions of the tools should be validated not only by country teams but also by the organization, and encompassing both linguistic and psychometric validation.

Funding: The study received no external funding.

Competing Interests: The authors have no relevant financial or non-financial interests to disclose.

Author Contributions: SM: conceptualization, methodology, formal analysis, writing – original draft; MBP: writing – original draft, review & editing; ML: writing – review & editing, contribution to data collection; AJ: writing – review & editing, contribution to data collection. All authors read and approved the final manuscript.

Ethics approval: No ethical approval is required for this study.

Acknowledgements: The authors would like to thank everyone who assisted in the analysis of the tools: first of all Göran Gustafson for his assistance in recruiting volunteers; and the native-speakers volunteers: Artur Hamlin, Veronika Jaeger, Hans Meerveld, Eleferios Meletis, Polychronis Kostoulas, Anna-Liisa Puttonen, and Marli Zambrano.

The authors would like to also express their gratitude to Prof. Mateusz Jankowski for his valuable support and guidance during the analysis.

References

1. S.M. Skevington, M. Lotfy, K.A. O'Connell, The World Health Organization's WHOQOL-BREF quality of life assessment: Psychometric properties and results of the international field trial. A Report from the WHOQOL Group, *Quality of Life Research* 13 (2004) 299–310. <https://doi.org/10.1023/B:QURE.0000018486.91360.00>.
2. WHOQOL Group, Study protocol for the World Health Organization project to develop a Quality of Life assessment instrument (WHOQOL), *Quality of Life Research* 2 (1993) 153–159. <https://doi.org/10.1007/BF00435734>.

3. L. Cid, D. Monteiro, D.S. Teixeira, A. Evmenenko, A. Andrade, T. Bento, A. Vitorino, N. Couto, F. Rodrigues, Assessment in Sport and Exercise Psychology: Considerations and Recommendations for Translation and Validation of Questionnaires, *Front Psychol* 13 (2022) 806176. <https://doi.org/10.3389/fpsyg.2022.806176>.
4. J.M. Ramada-Rodilla, C. Serra-Pujadas, G.L. Delclós-Clanchet, Adaptación cultural y validación de cuestionarios de salud: revision y recomendaciones metodológicas, *Salud Publica Mex* 55 (2013) 57–66. <https://www.scielosp.org/pdf/spm/2013.v55n1/57-66/es>.
5. S.L. Eremenco, D. Cella, B.J. Arnold, A comprehensive method for the translation and cross validation of health status questionnaires, *Eval Health Prof* 28 (2005) 212–232. <https://doi.org/10.1177/0163278705275342>.
6. R.K. Hambleton, The next generation of the ITC test translation and adaptation guidelines, *European Journal of Psychological Assessment* 17 (2001) 164–172. <https://doi.org/10.1027/1015-5759.17.3.164>.
7. F. Vijver, R. Hambleton, Translating tests. some practical guidelines, *Eur. Psychol.* 1 (1996) 89–99. <https://doi.org/10.1027/1016-9040.1.2.89>.
8. C.T. Beck, H. Bernal, R.D. Froman, Methods to document semantic equivalence of a translated scale, *Res Nurs Health* 26 (2003) 64–73. <https://doi.org/10.1002/nur.10066>.
9. S.P. McKenna, L.C. Doward, The translation and cultural adaptation of patient-reported outcome measures, *Value Health* 8 (2005) 89–91. <https://doi.org/10.1111/j.1524-4733.2005.08203.x>.
10. D. Wild, A. Grove, M. Martin, S. Eremenco, S. McElroy, A. Verjee-Lorenz, others, Principles of Good Practice for the Translation and Cultural Adaptation Process for Patient-Reported Outcomes (PRO) Measures: report of the ISPOR Task Force for Translation and Cultural Adaptation, *Value Health* 8 (2005) 94–104. <https://doi.org/10.1111/j.1524-4733.2005.04054.x>.
11. B.D. Zumbo, E.K. Chan, *Validity and validation in social, behavioral, and health sciences*, Springer International Publishing, Switzerland, 2014.
12. American Psychological Association, *Standards for educational and psychological testing*, American Educational Research Association, Washington, DC, 2014.
13. M. Herdman, J. Fox-Rushby, X. Badia, 'Equivalence' and the translation and adaptation of health-related quality of life questionnaires, *Qual Life Res.* 6 (1997) 237–247. <https://doi.org/10.1023/A:1026410721664>.
14. M.R. Mullen, Diagnosing measurement equivalence in cross-national research, *J Int Bus Stud.* 26 (1995) 573–596. <https://doi.org/10.1057/palgrave.jibs.8490187>.
15. D. Wild, S. Eremenco, I. Mear, M. Martin, C. Houchin, M. Gawlicki, A. Hareendran, I. Wiklund, L.Y. Chong, R. von Maltzahn, others, Multinational trials—recommendations on the translations required, approaches to using the same language in different countries, and the approaches to support pooling the data: the ISPOR Patient-Reported Outcomes Translation and Linguistic Validation Good Research Practices Task Force report, *Value Health* 12 (2009) 430–440. <https://doi.org/10.1111/j.1524-4733.2008.00471.x>.
16. D.E. Beaton, C. Bombardier, F. Guillemin, M.B. Ferraz, Guidelines for the process of cross-cultural adaptation of self-report measures, *Spine (Phila Pa 1976)* 25 (2000) 3186–3191. <https://doi.org/10.1097/00007632-200012150-00014>.
17. B.M. Byrne, T.L. Campbell, Cross-cultural comparisons and the presumption of equivalent measurement and theoretical structure a look beneath the surface, *J Cross-Cult Psychol.* 30 (1999) 555–574. <https://doi.org/10.1177/0022022199030005001>.
18. B.M. Byrne, D. Watkins, The issue of measurement invariance revisited, *J Cross-Cult Psychol.* 34 (2003) 155–175. <https://doi.org/10.1177/0022022102250225>.
19. J. Epstein, R.M. Santo, F. Guillemin, A review of guidelines for cross-cultural adaptation of questionnaires could not bring out a consensus, *J Clin Epidemiol* 68 (2015) 435–441. <https://doi.org/10.1016/j.jclinepi.2014.11.021>.
20. J.A. Flaherty, others, Developing instruments for cross-cultural psychiatric research, *J Nerv Ment Dis.* 176 (1988) 257–263. <https://doi.org/10.1097/00005053-198805000-00001>.
21. N. Luo, others, Do English and Chinese EQ-5D versions demonstrate measurement equivalence? An exploratory study, *Health Qual Life Outcomes.* 1 (2003) 1. <https://link.springer.com/doi/10.1186/1477-7525-1-7>.

22. M.E. Oliveri, R. Lawless, J.W. Young, A validity framework for the use and development of exported assessments, Princeton, NJ, 2015.
23. C. Acquadro, B. Jambon, D. Ellis, P. Marquis, Language and translation issues, in: B. Spilker (Ed.), *Quality of Life and Pharmacoeconomics in Clinical Trials*, 2nd Ed., Lippincott-Raven, Philadelphia, PA, 1996: pp. 575–585.
24. I.B. Corless, P.K. Nicholas, K.M. Nokes, Issues in cross-cultural quality of life research, *Journal of Nursing Scholarship* 33 (2001) 15–20. <https://doi.org/10.1111/j.1547-5069.2001.00015.x>.
25. W.R. Lenderking, Comments on the ISPOR Task Force Report on Translation and Adaptation of Outcomes Measures: guidelines and the need for more research, *Value Health* 8 (2005) 92–93. <https://doi.org/10.1111/j.1524-4733.2005.08201.x>.
26. B.M. Van Widenfelt, P.D. Treffers, E. de Beurs, B.M. Siebelink, E. Koudijs, Translation and cross-cultural adaptation of assessment instruments used in psychological research with children and families, *Clin Child Fam Psychol Rev* 8 (2005) 135–147. <https://doi.org/10.1007/s10567-005-4752-1>.
27. A. Kjaergaard Danielsen, H.C. Pommergaard, J. Burcharth, E. Angenete, J. Rosenberg, Translation of Questionnaires Measuring Health Related Quality of Life Is Not Standardized: A Literature Based Research Study, *PLoS One* 10 (2015) e0127050. <https://doi.org/10.1371/journal.pone.0127050>.
28. O. Chassany, P. Sagnier, P. Marquis, S. Fulleton, N. Aaronson, Patient reported outcomes and regulatory issues: the example of health-related quality of life - a European guidance document for the improved integration of HRQL assessment in the drug regulatory process, *Drug Inform J* 36 (2002) 209–238. <https://doi.org/10.1177/009286150203600127>.
29. U.S. Department of Health, H. Services, Guidance for Industry: Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims. Final Guidance., (2009).
30. W. Maneesriwongul, J. Dixon, Instrument translation process: A methods review, *J Adv Nurs* 48 (2004) 175–186. <https://doi.org/10.1111/j.1365-2648.2004.03185.x>.
31. M. Hawkins, C. Cheng, G.R. Elsworth, R. Osborne, Translation method is validity evidence for construct equivalence: analysis of secondary data routinely collected during translations of the Health Literacy Questionnaire (HLQ), *BMC Med Res Methodol* 20 (2020) 130. <https://doi.org/10.1186/s12874-020-00962-8>.
32. C. Acquadro, others, Emerging good practices for translatability assessment (TA) of patient-reported outcome (PRO) measures, 2 (2018) ssss.
33. A.M. Chang, J.P.C. Chau, E. Holroyd, Translation of questionnaires and issues of equivalence, *J Adv Nurs* 29 (1999) 316–322. <https://doi.org/10.1046/j.1365-2648.1999.00891.x>.
34. M. Hawkins, G.R. Elsworth, R.H. Osborne, Application of validity theory and methodology to patient-reported outcome measures (PROMs): building an argument for validity, *Qual Life Res.* (2018) 1–16.
35. C. Acquadro, K. Conway, A. Hareendran, N. Aaronson, Literature review of methods to translate health-related quality of life questionnaires for use in multinational clinical trials, *Value Health* 11 (2004) 509–521. <https://doi.org/10.1111/j.1524-4733.2007.00292.x>.
36. M. Gonçalves, M. Simões, L. Almeida, C. Machado, Avaliação psicológica. instrumentos validados para a população Portuguesa, in: M. Gonçalves, M. Simões, L. Almeida, C. Machado (Eds.), [Psychological Assessment. Validated Measures to Portuguese Population], 2nd Edn, Vol. I, Quarteto, Coimbra, 2006.
37. F. Guillemin, C. Bombardier, D. Beaton, Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines, *J Clin Epidemiol* 46 (1993) 1417–1432. [https://doi.org/10.1016/0895-4356\(93\)90142-N](https://doi.org/10.1016/0895-4356(93)90142-N).
38. M. Koller, N.K. Aaronson, J. Blazeby, A. Bottomley, L. Dewolf, P. Fayers, others, Translation procedures for standardised quality of life questionnaires: The European Organisation for Research and Treatment of Cancer (EORTC) approach, *Eur J Cancer* 43 (2007) 1810–1820. <https://doi.org/10.1016/j.ejca.2007.05.029>.
39. K. Oj, J. Mainz, A. Helin, M. Ribacke, F. Olesen, P. Hjortdahl, Oversettelse av spørreskjema. Et oversett metodeproblem., *Nord Med* 113 (1998) 363–366.
40. A.D. Sperber, Translation and validation of study instruments for cross-cultural research, *Gastroenterology* 126 (2004) S124–S128. <https://doi.org/10.1053/J.GASTRO.2003.10.016>.
41. C.T. Su, L.D. Parham, Generating a valid questionnaire translation for cross-cultural use, *Am J Occup Ther* 56 (2002) 581–585. <https://doi.org/10.5014/AJOT.56.5.581>.

42. S. Tsang, C.F. Royse, A.S. Terkawi, Guidelines for developing, translating, and validating a questionnaire in perioperative and pain medicine, *Saudi J Anaesth* 11 (2017) S80–S89. https://doi.org/10.4103/sja.SJA_203_17.
43. R. Vallerand, Vers une méthodologie de validation transculturelle de questionnaires psychologiques: implications pour la recherche en langue française [Toward a methodology for the transcultural validation of psychological questionnaires: implications for research in the French language], *Can. Psychol.* 30 (1989) 662–680. <https://doi.org/10.1037/h0079856>.
44. D. Grifee, Questionnaire translation and questionnaire validation: Are they the same?, in: Paper Presented at the Annual Meeting of the American Association for Applied Linguistics, St. Louis, MO, 2001.
45. S. Maksymowicz, P. Kukołowicz, T. Siwek, A. Rakowska, Validation of the revised Amyotrophic Lateral Sclerosis Functional Rating Scale in Poland and its reliability in conditions of the medical experiment, *Neurological Sciences* (2020). <https://doi.org/10.1007/s10072-020-04565-5>.
46. A. Anum, S. Adjorlolo, C. Akotia, A. Aikins, Validation of the multidimensional WHOQOL-OLD in Ghana: A study among population-based healthy adults in three ethnically different districts, *Brain Behav* 11 (2021). <https://doi.org/10.1002/brb3.2193>.
47. A. Muhye, N. Fentahun, Validation of Quality-of-Life assessment tool for Ethiopian old age people, *F1000Res* 12 (2024) 282. <https://doi.org/10.12688/f1000research.130379.2>.
48. A. Ahmed, M. Saqlain, N. Akhtar, F. Hashmi, A. Blebil, J. Dujaili, M. Umair, A. Bukhsh, Translation and cross-cultural adaptation of WHOQOL-HIV Bref among people living with HIV/AIDS in Pakistan, *Health Qual Life Outcomes* 19 (2021). <https://doi.org/10.1186/s12955-021-01693-0>.
49. A.O. Akinpelu, F.A. Maruf, B.O. Adegoke, Validation of a Yoruba translation of the World Health Organization’s quality of life scale–short form among stroke survivors in Southwest Nigeria, *Afr J Med Med Sci* 35 (2006) 417–424.
50. T. Gebrye, E. Okoye, C. Akosile, I. Onwuakagba, R. Uwakwe, C. Igweze, V. Chukwuma, F. Fatoye, Adaptation and Validation of the Nigerian (Igbo) Version of the WHOQOL-OLD Module, *Res Soc Work Pract* 33 (2022) 798–809. <https://doi.org/10.1177/10497315221130527>.
51. K. Lee, H. Jang, H. Choi, Korean translation and validation of the WHOQOL-DIS for people with spinal cord injury and stroke., *Disabil Health J* 10 (2017) 627–631. <https://doi.org/10.1016/j.dhjo.2016.12.017>.
52. S. Nedjat, A. Montazeri, K. Holakouie, K. Mohammad, R. Majdzadeh, Psychometric properties of the Iranian interview-administered version of the World Health Organization’s Quality of Life Questionnaire (WHOQOL-BREF): A population-based study, *BMC Health Serv Res* 8 (2008) 61. <https://doi.org/10.1186/1472-6963-8-61>.
53. H. Rezaeipandari, M. Morowatisharifabad, A. Mohammadpoorasl, A. Shaghghi, Cross-cultural adaptation and psychometric validation of the World Health Organization quality of life-old module (WHOQOL-OLD) for Persian-speaking populations, *Health Qual Life Outcomes* 18 (2020). <https://doi.org/10.1186/s12955-020-01316-0>.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.