

Journal Pre-proof

Psychometric properties of the experimental and modified EQ-HWB and EQ-HWB-9 in patients with anxiety and/or depression in China

Richard Huan Xu, PhD, Chenxi Yang, MSc, Yuanshuo Xu, BSc, Ole Marten, PhD, Fanni Rencz, DSc

PII: S1098-3015(26)02521-0

DOI: <https://doi.org/10.1016/j.jval.2026.06.004>

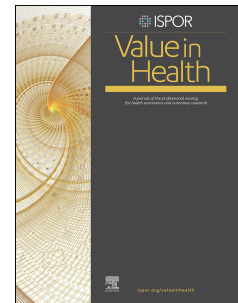
Reference: JVAL 4708

To appear in: *Value in Health*

Received Date: 11 December 2025

Revised Date: 2 June 2026

Accepted Date: 4 June 2026



Please cite this article as: Xu RH, Yang C, Xu Y, Marten O, Rencz F, Psychometric properties of the experimental and modified EQ-HWB and EQ-HWB-9 in patients with anxiety and/or depression in China, *Value in Health* (2026), doi: <https://doi.org/10.1016/j.jval.2026.06.004>.

This is a PDF of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability. This version will undergo additional copyediting, typesetting and review before it is published in its final form. As such, this version is no longer the Accepted Manuscript, but it is not yet the definitive Version of Record; we are providing this early version to give early visibility of the article. Please note that Elsevier's sharing policy for the Published Journal Article applies to this version, see: <https://www.elsevier.com/about/policies-and-standards/sharing#4-published-journal-article>. Please also note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Copyright © 2026, International Society for Pharmacoeconomics and Outcomes Research, Inc.
Published by Elsevier Inc.

Full title: Psychometric properties of the experimental and modified EQ-HWB and EQ-HWB-9 in patients with anxiety and/or depression in China

Authors: Richard Huan Xu¹ PhD, Chenxi Yang¹ MSc, Yuanshuo Xu¹ BSc, Ole Marten² PhD, Fanni Rencz^{3,4} DSc

Affiliations:

1. Department of Rehabilitation Sciences, Faculty of Health and Social Sciences, Hong Kong Polytechnic University, Hong Kong SAR, China
2. School of Public Health, University of Bielefeld, Bielefeld, Germany
3. Department of Health Policy, Corvinus University of Budapest, Budapest, Hungary
4. EuroQol Research Foundation, Rotterdam, The Netherlands

Correspondence author: RX, richard.xu@polyu.edu.hk

Running title: EQ-HWB in patients with anxiety/depression

Precis: The study examines the psychometric properties of the EQ-HWB in a clinical population of individuals with anxiety and/or depression in China.

Word count: 4271

Number of pages: 37

Number of figures: 2

Number of tables: 4

Number of supplementary materials: 4

Funding: This study is funded by EuroQol Research Foundation (ID:1897-RA).

Psychometric properties of the experimental and modified EQ-HWB and EQ-HWB-9 in patients with anxiety and/or depression in China

Abstract

Objectives: We aimed to examine the psychometric properties of the EuroQol Health and Wellbeing (EQ-HWB) and its short form (EQ-HWB-9) in patients with anxiety and/or depression. A secondary objective was to assess the performance of item modifications proposed by the EQ-HWB developers after the initial version was published.

Methods: A multicentre cross-sectional study included 900 patients with anxiety and/or depression in China. The survey included the modified 25-item EQ-HWB and the 12 experimental items that differ between the experimental (2022) and modified (2024) versions. EQ-HWB-9 responses were derived from the EQ-HWB (*modified version, 2024*). Psychometric evaluation included distributional characteristics, convergent validity (correlations with EQ-5D-5L, GAD-7, PHQ-9, ReQoL-10, SWEMWBS), known-groups validity, test-retest reliability and exploratory cross-instrument dimensionality (principal component analysis).

Results: No ceiling effect was observed (0.8% EQ-HWB, 2.4% EQ-HWB-9). The sleep (9.0%), anxiety (9.4%), and exhaustion (13.4%) items had the lowest ceilings. Convergent validity was strong, particularly with PHQ-9 and GAD-7 (items ≤ 0.78 ; scale ≤ 0.90). Known-groups validity showed large effect sizes across anxiety and depression severity groups. Test-retest reliability was good (item-level Gwet's AC2 0.60–0.97; instrument-level ICC 0.72–0.73). EQ-

HWB items loaded on five factors (mental health and cognition, positive wellbeing, physical health, pain and discomfort, social relationships). The differences in psychometric performance between the modified and experimental items varied across items.

Conclusion: The EQ-HWB/EQ-HWB-9 demonstrate good validity and reliability for assessing health and wellbeing in Chinese patients with anxiety and/or depression. Results support the potential usefulness of the two instruments for cross-sectoral decision-making in anxiety and/or depression.

Keywords: EQ-HWB; anxiety; depression; mental health; psychometrics

Highlights

- Psychometric properties of the modified EQ-HWB have not previously been evaluated in patients with common mental health conditions, and no quantitative studies have directly compared the impact of item modifications on psychometric performance against the experimental version.
- In patients with anxiety and/or depression, the modified EQ-HWB and EQ-HWB-9 showed no ceiling or floor effects, good convergent validity, known-groups validity by severity, and acceptable test–retest reliability. Differences in performance between experimental and modified items were generally small, with the largest discrepancies for positively versus negatively framed items.
- Overall, our findings support the validity of the modified EQ-HWB and EQ-HWB-9 for assessing health and wellbeing in mental health populations and for informing cross-sector (health and social care) evaluations and decision-making.

Introduction

Globally, approximately 5.7% and 4.4% of adults suffer from depression and anxiety disorders, respectively.^{1,2} In China, an estimated 54 million people suffer from depression, and 41 million have anxiety disorders.³ These two conditions frequently co-occur. Around 50–85% of people with major depression also experience clinically significant anxiety symptoms or a comorbid anxiety disorder, and up to 90% of patients with anxiety disorders show depressive symptoms or comorbid depression.⁴ Anxiety and depression significantly impair psychosocial functioning,⁵ causing substantial deficits in physical, emotional, and social domains that reduce the ability to perform daily activities and enjoy life.⁶ These disorders also tend to diminish overall quality of life, including life satisfaction, happiness, fulfillment, sense of purpose, and achievement.⁷

The EQ-HWB (EuroQol Health and Wellbeing) is a generic instrument designed to measure health and wellbeing, encompassing aspects of quality of life that extend beyond traditional health-related domains, including those relevant to social care, caregivers and broader wellbeing.⁸ The EQ-HWB currently has experimental status and remains under development. The EQ-HWB has two versions (long and short), designed for use across health care, social care, and public health.⁹ The long form is a broader profile measure, while the short version, EQ-HWB-9, includes a subset of the items from the long form. Preference weights may be attached to the EQ-HWB-9, making it potentially suitable for use in certain economic evaluations, particularly in cross-sectoral applications (e.g., health care and social care).⁸

The growing body of literature on the psychometric properties of the EQ-HWB and EQ-HWB-9 demonstrates promising psychometric performance, primarily in general population, caregiver, and mixed samples. Studies across different countries have shown strong construct validity, convergent validity with measures of health and mental well-being scales, acceptable test-retest reliability, no or low floor/ceiling effects in relevant domains, and good known-group discrimination.¹⁰⁻²⁷ Despite these encouraging findings, most evidence to date comes from non-clinical or heterogeneous samples, and there remains a gap in focused evaluation among people with specific mental health conditions in clinical contexts.^{12,13,17,27} Testing the psychometric properties of the EQ-HWB and EQ-HWB-9 in patients with anxiety or depression is crucial because this instrument was specifically designed to capture a broader range of health and wellbeing impacts, including key domains such as hopelessness, loneliness, and control, that are central to mental disorders.

Further psychometric testing is essential because the instrument remains experimental, and recent modifications, including changing three positively framed items to negative, changing certain response options in the frequency and difficulty response scales, and item wording changes (e.g., replacing 'and' with 'or' in the mobility item) were proposed by the developer team. These changes require targeted re-validation in priority clinical groups with mental health conditions to confirm that they enhance (or at least preserve) reliability, validity, and sensitivity.²⁸

Thus, the primary objective was to examine the psychometric properties of the EQ-HWB and EQ-HWB-9 in Chinese patients with anxiety and/or depression. The secondary objective

was to compare the recently proposed item modifications against the items in the original experimental version ²⁹ to inform the refinement of the instrument.

Methods

Study design and recruitment

We conducted a multi-centre cross-sectional survey in China between January and September 2025. Adult patients with anxiety and/or depression were recruited from three collaborative hospitals in Guangdong and Zhejiang provinces. A convenience sampling method was employed to recruit participants in accordance with our inclusion and exclusion criteria. Recruitment and screening of patients were conducted in the outpatient department by clinical psychologists. The inclusion criteria were as follows: a) individuals aged ≥ 18 years; b) able to read and speak Chinese; c) no cognitive problems (assessed by a clinical psychologist using the MoCA, >26);³⁰ and d) able to provide informed consent. Individuals with severe mental health issues, such as severe psychosis or cognitive impairment, were excluded to ensure accurate self-reporting. Participants with severe physical health conditions that could interfere with their ability to complete the survey were deemed ineligible. Our aim was to recruit all patients with mild, moderate, and severe anxiety and/or depression; however, final participation was determined based on the psychologist's professional judgment in clinical practice.

For patients meeting inclusion criteria, psychologists briefly introduced the study and assessed willingness to participate. Consenting patients were referred to the study team. A research assistant then contacted them to provide detailed study information, obtained written informed consent, and guided them to access the online questionnaire by scanning

a WeChat QR code. All assistants received standardized training to ensure assistance was consistent and non-directive. Missing responses were permitted so that the feasibility of the instrument could be evaluated. The questionnaire comprised several sections in a fixed order and took around 10 minutes to complete. The first section was the modified version of the EQ-HWB, followed by the EQ-5D-5L in the second section. The third one assessed mental health using the Generalized Anxiety Disorder Scale-7 (GAD-7) and Patient Health Questionnaire-9 (PHQ-9). The fourth section included the experimental version of the EQ-HWB (12 items only), while the fifth section included the Recovering Quality of Life-10 (ReQoL-10) and Short Warwick-Edinburgh Mental Wellbeing Scale (SWEMWBS). The final section collected sociodemographic information. All instruments were administered using their official validated Chinese versions. All participants were invited to complete the EQ-HWB modified and experimental (12 items only) versions again two weeks later to examine the test-retest reliability of the EQ-HWB.

To ensure data quality, several measures were implemented. The research staff was well-trained to administer all baseline assessments face-to-face using a tablet-based system with built-in range checks, skip logic, and mandatory response fields to minimize missing data and entry errors. Standardized scripts and protocols were used to maintain consistency. Additionally, attention and consistency checks across related instruments were embedded within the assessments. Protocol and informed consent were approved by the Institutional review board of the *Blind for review* (Ref ID: *Blind for review*).

Measures

EQ-HWB

The EQ-HWB was translated into Simplified Chinese under the supervision of the EuroQol Version Management Committee earlier and has been validated in the Chinese population.¹¹ In this study, participants self-completed the 25-item EQ-HWB (v1.2, modified 2024 version) along with 12 unmodified items from the 25-item experimental version (v1.1, 2022 version). The 25-item EQ-HWB experimental was not presented to respondents in full to reduce cognitive burden of the survey. Hereafter, we refer to these as 'modified' and 'experimental' in this manuscript. A summary of the differences between the two versions is presented in the Appendix (Table A1). Several instruments (e.g., PHQ-9) were used to separate the 25 and 12 items. All EQ-HWB items are rated on a five-point scale describing difficulty, frequency, or severity, with a 7-day recall period. Level Sum Scores (LSSs) were calculated for both the EQ-HWB and EQ-HWB-9 by summing the item-level responses (*1 = best outcome to 5 = worst outcome*), transformed to a 0–100 scale, and reversed so that higher scores indicated better outcomes.

EQ-5D-5L

The Simplified Chinese version of the EQ-5D-5L was administered.³¹ The EQ-5D-5L consists of five items, each rated on five levels, to assess an individual's HRQoL on the day of the survey. Its psychometric properties have been widely confirmed in various Chinese populations.^{32–37} LSSs were calculated for EQ-5D-5L (range: 5–25). The EQ VAS, which measures self-rated health on a scale from 0 to 100, was also completed. The EQ-5D-5L was included solely as a reference measure to assess the convergent validity and explore the cross-instrument dimensionality of the EQ-HWB and was not used for direct comparative performance analyses between the instruments.

PHQ-9

The validated Chinese version of the PHQ-9 was used to measure depressive symptoms over the past two weeks.³⁸ Each of its nine items is rated on a 4-point scale. The cut-off points of 0–4 indicate no depressive symptoms, 5–9 mild depressive symptoms, 10–14 moderate depressive symptoms, 15–19 moderately-severe depressive symptoms, and 20–27 severe depressive symptoms.³⁹

GAD-7

The validated Chinese version of the GAD-7 was used to measure individuals' anxiety levels over the past two weeks.⁴⁰ Each of its seven items is rated on a 4-point scale. The levels of anxiety were classified as follows: no or minimal (0–4), mild (5–9), moderate (10–14), and severe (15–21).⁴¹

ReQoL-10

The validated Chinese version of the ReQoL-10, a condition-specific preference-based instrument, was used to assess the quality of life for people with different mental health conditions.⁴² All items are scored on a five-level scale, ranging from 'None of the time' to 'Most of the time'. A summative score (0–40) based on 10 mental health-related items was calculated, where a higher score indicates better quality of life.

SWEMWBS

The validated Chinese version of the SWEMWBS is a 7-item instrument developed to assess mental wellbeing at the population level and to evaluate the effectiveness of projects and policies aimed at improving it.⁴³ The SWEMWBS is scored by summing the responses to its 7

items, each rated on a 5-point Likert scale from 1 to 5, resulting in a total score ranging from 7 to 35; higher scores indicate greater mental wellbeing.⁴⁴

Data analysis

All analyses were conducted at both the item and scale levels, where relevant. For the primary objective (psychometric assessment of the EQ-HWB and EQ-HWB-9), responses for the EQ-HWB-9 were extracted directly from the 25-item EQ-HWB (modified). For the secondary objective (assessment of item modifications), we compared the performance of the modified (2024) and experimental (2022) versions of the 12 differing items across all measurement properties.

Descriptive characteristics

The relative frequencies of responses for each item in the descriptive system of the EQ-HWB were computed. Bar charts were used to visually represent the distributions of item responses between the two versions. Ceiling and floor effects at both items and overall score levels were examined. Ceiling and floor effects were defined as the proportion of the sample endorsing the best and worst response option on an item, or the highest or lowest scores on the scale, respectively. A high ceiling or floor indicates limitations in the item's or measure's sensitivity. McNemar's test was used to assess whether there were significant differences in the proportion of respondents reporting ceiling and floor between the corresponding experimental and modified items. A threshold of 70% was used at the item level, following previous EQ-HWB validation studies,⁴⁵ whereas a 15% threshold was applied at the scale level.⁴⁶

Test-retest reliability

To simplify response options and improve interpretability, we revised the global health status scale by combining 'excellent', 'very good', and 'good' into a single 'good' category, resulting in a three-level scale of good/fair/poor;⁴⁷ only participants who reported no change in their health status at 2 weeks, as measured by this scale, were included in the assessment of test-retest reliability. Gwet's AC2 values were used to assess item-level test-retest reliability. These values were interpreted using the following cutoff points: <0 poor, 0.01–0.20 slight, 0.21–0.40 fair, 0.41–0.60 moderate, 0.61–0.80 substantial, 0.81–0.99 almost perfect or 1.00 perfect.⁴⁸ The scale-level test-retest reliability was examined using the intraclass correlation coefficient (ICC, two-way random effects, absolute agreement), with values ≥ 0.7 considered acceptable.⁴⁹

Informativity

Informativity was computed for each item in the descriptive systems of the EQ-HWB and the EQ-HWB-9 using the Shannon index (H') for absolute informativity and the Shannon evenness index (J') for relative informativity.⁵⁰ Calculations were performed separately for the modified and experimental items, respectively.

Exploratory cross-instrument dimensionality analysis

Principal component analysis (PCA), in line with the COSMIN guideline recommendations,²⁸ was conducted to examine the underlying dimensionality of the items from the EQ-HWB in combination with an item pool from all the other study instruments. This analysis was exploratory in nature. The results were interpreted as preliminary insights into the overall construct space rather than definitive evidence for the factor structure of the EQ-HWB

itself. Bartlett's test of sphericity ($p < 0.05$) and the Kaiser–Meyer–Olkin Test ($KMO > 0.6$) were conducted to verify the assumptions for PCA. The number of factors retained was determined using the eigenvalue-greater-than-1 criterion.⁵¹ Given the expected associations between factors, Oblimin rotation was applied. Principal Axis factoring was used for factor extraction, with a factor loading cut-off of ≥ 0.3 for item retention.⁵²

Convergent validity

Convergent validity was assessed through hypothesis testing (Appendix Table A2). We hypothesized that measures assessing conceptually similar constructs would exhibit at least moderate correlations with the EQ-HWB, both at the item level (between corresponding items) and at the scale level (between overall scores). For example, the EQ-HWB item "Sadness/depression" was expected to correlate with the "Anxiety/depression" item from the EQ-5D-5L, and the EQ-HWB item "Sleep" was expected to correlate with the PHQ-9 item "Trouble falling or staying asleep, or sleeping too much." These correlations were evaluated using Spearman's correlation coefficients and categorized as weak (< 0.3), moderate ($0.3-0.49$), or strong (≥ 0.5).⁵³ To enable targeted item-level assessment, we examined correlations between specific EQ-HWB items and the most conceptually overlapping single item from the other instruments, thereby clarifying which items of the EQ-HWB are most closely aligned with established constructs.^{13,18} We also assessed the correlations between the EQ-HWB/EQ-HWB-9 LSSs and the total scores of the other standardized instruments, using Pearson's correlation coefficient interpreted under the same criteria for Spearman's correlation coefficients.

Known-group validity

One-way ANOVA was used to assess whether the EQ-HWB discriminates between known groups (e.g., severity of depression/anxiety) at the item and scale levels. Cohen's d (for two-group comparison: ≥ 0.8 large, 0.5-0.79 medium, 0.2-0.49, small, and < 0.2 , trivial) and η^2 (for three-group comparison: ≥ 0.14 large, 0.06-0.13 medium, and 0.01-0.05, small) were calculated to reflect effect size.⁵⁴ We expected worse health and wellbeing (i.e., lower scores for EQ-HWB/EQ-HWB-9 LSS) to be associated with factors such as longer disease duration, poor general health (global health question), presences of co-morbidities (chronic and mental health conditions), psychiatric medication prescribed, and more severe levels of depression and/or anxiety.

Results

Patients' characteristics

A total of 900 patients completed the survey (response rate = 81%), with mean completion time of 10.7 minutes. Of these, 61.7% were female, and 48.3% had completed tertiary education or higher. The mean age was 37.7 years, and the mean time since first diagnosis of anxiety or depression was 4 years. Approximately 4.4% of participants reported severe anxiety on the GAD-7, and 11.5% reported severe depression on the PHQ-9. Details about the characteristics of the study population are presented in the Appendix (Table A3).

Distributional characteristics

One missing response was recorded for the EQ-HWB items 4, 8, and 24. The distributions of responses on all modified and experimental items are presented in Figure 1. The items "getting around inside and/or outside" and "personal care" showed the highest ceilings with 76.9% and 77.1%, respectively. The items "sleep" (9.0%), "anxiety" (9.4%), and "exhaustion"

(13.4%) had the lowest ceilings. Compared to the experimental version, the modified items resulted in a mixed pattern (Table 1). Among the 12 corresponding items, the ceilings for the modified versions were significantly higher in 4 cases, lower in 4 cases, and not significantly different in 4 cases, while the floors were lower in 3 cases, higher in 2 cases, and not different in 7 cases. The biggest differences in distribution occurred in the three positively framed items from the experimental version, which were rephrased to negative framing in the modified version: "Do things one wanted to do" became "Difficulty with enjoyable activities," with the ceiling rising from 12.9% to 53.8%; "Accepted by others" to "Excluded by others," with the ceiling rising from 18.8% to 39.2%; and "Feel good about self" to "Negative feelings about self," with the ceiling rising from 15.1% to 22.1%. The floor effects for these three items decreased. The ceiling decreased for both cognitive items, from 20.8% to 13.9% for remembering and from 22.9% to 16.4% for concentrating/thinking clearly, whereas the floor slightly increased for both items. At the instrument level, no ceiling or floor effects were identified for the modified EQ-HWB or the EQ-HWB-9.

Informativity

The modified and experimental versions of the EQ-HWB and EQ-HWB-9 showed similar J' (Appendix, Table 4). Among modified items, J' were generally comparable, although the experimental version performed slightly better for several items. The largest difference was observed for the item "Do things one wanted to do" ($H' = 2.28$ in the experimental version vs 1.72 in the modified version).

Test-retest reliability

Test–retest reliability was assessed in the 132 out of 228 participants who reported the same general health status at 2 weeks as at baseline. Item-level agreement, measured using Gwet’s AC2, ranged from 0.67 (exhaustion) to 0.97 (personal care, hearing, and getting around inside or outside). Of the 25 items from the modified EQ-HWB, 12 exhibited substantial agreement and 13 showed almost perfect agreement. Comparing the corresponding modified and experimental items, seven modified items had higher Gwet’s AC2 values and five had lower values. The most substantial improvements occurred among items that were reworded from positive to negative framing. At the scale level, the ICC for the modified EQ-HWB and EQ-HWB-9 was 0.77 and 0.73, respectively (Table 1), thus indicating acceptable test-retest reliability.

Exploratory cross-instrument dimensionality

The KMO measure of 0.98 and Bartlett’s test ($p < 0.001$) indicated excellent suitability for factor analysis. The PCA revealed a five-component structure (cumulative explained variance = 62%), with the modified EQ-HWB loading on all components. The five components were: Mental health and cognition (e.g., “anxiety,” “sadness/depression,” “frustration,” “negative feelings about self,” “lacked control over daily life” from EQ-HWB, and all items from GAD-7 and PHQ-9; loadings 0.50–0.80), Positive wellbeing (six items from ReQoL and all items from the SWEMWBS; loadings 0.63–0.78), Physical health (e.g., “getting around inside or outside,” “personal care,” “hearing” from EQ-HWB and mobility and self-care from EQ-5D-5L; loadings 0.43–0.75), Pain and discomfort (e.g., “pain [frequency],” “pain [severity]” from EQ-HWB and pain/discomfort from EQ-5D-5L; loadings 0.62–0.78), and Social relationships (e.g., “loneliness,” “unsupported,” “excluded by others” from EQ-HWB; loadings 0.42–0.55). Cross-loadings on factors such as “sleep” and “hopeless” indicate

interconnections between domains, affirming the EQ-HWB's multidimensional coverage (Table 2).

Convergent validity

At the item-level, the modified EQ-HWB demonstrated strong convergent validity with other instruments, particularly for mental health-related items (Table 3). For instance, the anxiety item showed high correlations with GAD-7 items such as "Feeling nervous, anxious, or on edge" (0.70), while the sadness/depression item showed high correlations with the PHQ-9 item "Feeling down, depressed, or hopeless" (0.75). Physical and pain-related items also exhibited strong associations, with pain (severity) linking to EQ-5D-5L pain/discomfort (0.68). Social and wellbeing items, such as loneliness with ReQoL "I felt lonely" (0.78) and "negative feelings about self" with SWEMWBS "I've been feeling useful" (-0.57), further supported validity. Compared to the experimental version, correlations were mixed: some decreased modestly (e.g., day-to-day activities), while items rephrased positive to negative framing in the modified EQ-HWB showed increases, such as "Difficulty with enjoyable activities." At the scale-level, the modified EQ-HWB LSS and EQ-HWB-9 LSS displayed excellent convergent validity, with strong positive correlations with GAD-7 and PHQ-9.

Known-group validity

The modified EQ-HWB demonstrated significant discriminative ability across all seven groups, with large effect sizes in six groups, including anxiety severity, depression severity, comorbidities, and medication use (Table 4). The mean LSSs decreased significantly as anxiety and depression worsened, confirming the instrument's sensitivity to mental health

domains. For the EQ-HWB-9, the modified version performed comparably to the long form, with slightly larger effect sizes across most groups (4 groups versus 2 groups).

Figure 2 presents the known-group validity of the 12 corresponding modified and experimental EQ-HWB items by comparing effect sizes across a range of known groups. Overall, the modified items demonstrated comparable or slightly stronger discriminative ability than the experimental items in most groups. For example, for PHQ-9 group comparisons, the effect sizes were 0.01 (experimental) vs. 0.34 (modified) for “Accepted/Excluded by others,” 0.03 vs. 0.53 for “Feel good about self/Negative feelings about self,” and 0.03 vs. 0.26 for “Do things you want to do/Difficulty with enjoyable activities.”

Discussion

This study provides the first empirical evaluation of the psychometric properties of the modified EQ-HWB and EQ-HWB-9 in a clinical sample of patients with anxiety and/or depression. Both instruments demonstrated overall good performance, confirming their potential usefulness in mental health populations.

Ceiling effects were observed in some items, particularly in items “getting around inside and/or outside” and “personal care,” consistent with previous EQ-HWB studies.^{11,12,21,55,56} However, the ceilings were below the conventional threshold for most items, indicating sufficient sensitivity of the modified EQ-HWB in this population. Test–retest reliability was good, in line with previous studies in patient and caregiver populations using different versions of EQ-HWB, all of which fall within accepted “good” thresholds.^{15,20,21} Test-retest

reliability was comparable to the literature, with highest stability observed for physical-function items.

Convergent validity was robust at both item and scale levels, with strong correlations between EQ-HWB mental health items and corresponding GAD-7, PHQ-9, and ReQoL items. Scale-level correlations were excellent, indicating alignment with established measures of depression, anxiety, and mental wellbeing.¹³ Known-group validity further supported the instruments' discriminative power, with large effect sizes across most mental health-related predefined groups, as hypothesized. These findings demonstrate sensitivity to clinically meaningful differences and are consistent with results from previous studies using an earlier version of the EQ-HWB.^{12,13}

PCA results indicated that EQ-HWB items spread across five factors, a structure that differs from a prior four-factor model identified in a general Chinese population,¹¹ however, both affirm EQ-HWB's broad coverage, which seems to be an excellent fit for mental health populations. Although the positive wellbeing and social relationships factors include fewer items, they remain represented, for example, by the hopeless item loading on positive wellbeing and the unsupported and exclusion by other items on social relationships.

The revision of three EQ-HWB items from positive framing in the experimental to negative framing in the modified version represents a strategic, evidence-driven enhancement that may balance shifts in response patterns with substantial psychometric gains. This change seems to enhance test-retest reliability and strengthen discriminatory power, enabling better differentiation across predefined groups of patients.

Although previous studies have shown that generic HRQoL measures (e.g., the EQ-5D) perform adequately in mental health populations, the EQ-HWB was developed for populations, settings and contexts where the measurement of wellbeing in addition to health can inform the decision-making. Rather than replacing the EQ-5D, it serves a different purpose by allowing to measure broader aspects (e.g., loneliness, control) that may be particularly relevant for patients with anxiety and/or depression. There may be a potential value in the EQ-HWB-9 for interventions that influence wellbeing, such as psychological therapies (e.g., cognitive behavioral therapy), community-based mental health support programs and self-management interventions. The future use of the measure will largely depend on the country-specific context and on whether decision-makers are interested in considering these additional wellbeing aspects alongside health as measured by the EQ-5D-5L. EQ-HWB-9 would require preference weights to be used in such decision-making contexts. The EuroQol Group is currently developing a valuation protocol, including an exploration of appropriate methods for valuing this instrument.

This study has several strengths. It included a large, multi-site clinical sample and used a comprehensive set of outcome measures assessing symptom severity, quality of life, and positive mental wellbeing. This multi-domain approach provides convergent evidence of the EQ-HWB and EQ-HWB-9's performance across key constructs relevant to anxiety and depression. By demonstrating strong and consistent associations across these different measures, the findings enhance the robustness of the results within the sample. The results may be potentially generalizable to broader mental health populations (e.g., bipolar disorder, schizophrenia, obsessive-compulsive disorder) where anxiety or depression are

often among leading symptoms. However, dedicated validation studies in these populations are still needed to confirm the findings. Furthermore, our findings will also inform the ongoing refinement of EQ-HWB and support the future adoption of several proposed modifications.

Several limitations should be addressed as well. First, the order of administration for all instruments was fixed rather than randomized or counterbalanced across participants. Potential order effects and fatigue effects may have biased responses and affected observed correlations between measures. Especially, the 12 experimental items were consistently administered after the 25-item modified EQ-HWB version, which may introduce bias due to participants' prior exposure to the modified items, potentially influencing their responses to the experimental ones. Second, only 12 items from the experimental EQ-HWB were directly administered, which prevented comparisons between the experimental and modified EQ-HWB at the scale level. Third, test-retest reliability was assessed using unchanged health status to identify stable patients, focusing primarily on health aspects while potentially missing broader wellbeing components measured by the EQ-HWB. Additionally, the low follow-up rate and high stability in health status did not allow testing responsiveness to change. Moreover, we did not collect detailed information on medications or other interventions during the retest interval, so we could not fully account for potential treatment-related changes that might have influenced test-retest reliability estimates. Validity results between the EQ-HWB and other instruments involving differing recall periods and contextual framing may also be interpreted with caution. Furthermore, the sample's demographic skew toward younger, educated participants might not fully capture cultural or socioeconomic variations in mental health reporting within China.

Finally, we did not collect data on break-off rates or partial responses due to limited staffing in busy clinic settings; therefore, we could not formally assess respondent burden or fatigue. This may have influenced data quality in this mental health population.

Conclusion

The EQ-HWB and EQ-HWB-9 are reliable and valid instruments for the measurement of health and wellbeing in patients with anxiety and/or depression. The psychometric performance of the modified and experimental items varies by item, with the largest differences observed for those that shifted from positive to negative framing. These results support the potential usefulness of the two instruments for cross-sectoral decision-making among patients with anxiety and/or depression, which may extend to broader mental health populations with additional research.

References

1. Alonso J, Liu Z, Evans-Lacko S, et al. Treatment gap for anxiety disorders is global: Results of the World Mental Health Surveys in 21 countries. *Depress Anxiety*. 2018;35(3). doi:10.1002/da.22711
2. Moreno-Agostino D, Wu YT, Daskalopoulou C, Hasan MT, Huisman M, Prina M. Global trends in the prevalence and incidence of depression: a systematic review and meta-analysis. *J Affect Disord*. 2021;281. doi:10.1016/j.jad.2020.12.035
3. World Health Organization. Mental health - China. Accessed November 19, 2025. <https://www.who.int/china/health-topics/mental-health>
4. Steffen A, Nübel J, Jacobi F, Bätzing J, Holstiege J. Mental and somatic comorbidity of depression: A comprehensive cross-sectional analysis of 202 diagnosis groups using German nationwide ambulatory claims data. *BMC Psychiatry*. 2020;20(1). doi:10.1186/s12888-020-02546-8
5. Saris IMJ, Aghajani M, van der Werff SJA, van der Wee NJA, Penninx BWJH. Social functioning in patients with depressive and anxiety disorders. *Acta Psychiatr Scand*. 2017;136(4). doi:10.1111/acps.12774
6. Olesen J, Gustavsson A, Svensson M, Wittchen HU, Jönsson B. The economic cost of brain disorders in Europe. *Eur J Neurol*. 2012;19(1). doi:10.1111/j.1468-1331.2011.03590.x

7. Diener E, Suh EM, Lucas RE, Smith HL. Subjective well-being: Three decades of progress. *Psychol Bull.* 1999;125(2). doi:10.1037/0033-2909.125.2.276
8. Brazier J, Peasgood T, Mukuria C, et al. The EQ-HWB: Overview of the Development of a Measure of Health and Wellbeing and Key Results. *Value in Health.* 2022;25(4). doi:10.1016/j.jval.2022.01.009
9. Peasgood T, Mukuria C, Carlton J, et al. What is the best approach to adopt for identifying the domains for a new measure of health, social care and carer-related quality of life to measure quality-adjusted life years? Application to the development of the EQ-HWB? *European Journal of Health Economics.* 2021;22(7). doi:10.1007/s10198-021-01306-z
10. Lee P, Engel L, Lubetkin E, Gao L. Exploring the Comparability Between EQ-5D and the EQ Health and Wellbeing in the General Australian Population. *Value in Health.* Published online 2024. doi:10.1016/j.jval.2024.01.004
11. Xu RH, Rencz F. Psychometric Properties of the Modified EQ Health and Wellbeing and its Shorter Version in a General Chinese Population: A Comparative Analysis With EQ-5D-5L, SF-6Dv2, ICEpop CAPability measure for Adults, Short Version of Warwick-Edinburgh Mental Wellbeing Scale, and World Health Organization Well-Being Index 5-Item. *Value in Health.* 2025;28(8):1241-1252. doi:10.1016/j.jval.2025.04.2151
12. Zhang G, Yang Z, Luo N, Busschbach J. Comparing the Measurement Properties of the Preliminary Version of the EuroQol Health and Well-Being and EQ-5D-5L in Patients, Healthy General Public, and Caregivers. *Value Health Reg Issues.* 2025;47:101081. doi:10.1016/j
13. Keetharuth AD, Mukuria C, Peasgood T, Wailoo A. EQ Health and Wellbeing EQ-HWB: A Psychometric Assessment Across 6 Conditions and the General Population in the United Kingdom. *Value in Health.* Published online 2025. doi:10.1016/j.jval.2025.07.028
14. Bailey C, Dalziel K, Constable L, et al. The performance of the EQ-HWB-S as a measure of quality-of-life of caregivers in families that have experienced adverse events. *European Journal of Health Economics.* Published online 2024. doi:10.1007/s10198-024-01688-w
15. Pangestu S, Purba FD, Setyowibowo H, Azhar Y, Mukuria C, Rencz F. The Psychometric Properties of the EQ-HWB and EQ-HWB-S in Patients With Breast Cancer: A Comparative Analysis With EQ-5D-5L, FACT-8D, and SWEMWBS. *Value in Health.* Published online 2024. doi:https://doi.org/10.1016/j.jval.2024.12.003
16. Zhou W, Ding B, Busschbach J, Herdman M, Yang Z, Lu Y. EQ-5D-5L or EQ-HWB-S: Which is the Better Instrument for Capturing Spillover Effects in Parental Carers of Children with COVID-19? *Pharmacoeconomics.* 2025;43(5). doi:10.1007/s40273-025-01473-4
17. Bailey C, Dalziel K, Constable L, Peasgood T. Validity and Reliability of the EQ-HWB in Parents of Children With Health Conditions. *Value in Health.* 2025;28(9). doi:10.1016/j.jval.2025.05.013
18. Bailey C, Trapani K, Davies JN, Van Dam N, Galante J, Peasgood T. The psychometric performance of the EQ-HWB-9 for measuring health and wellbeing in a general population sample from Australia and New Zealand. *Quality of Life Research.* 2025;34(12). doi:10.1007/s11136-025-04061-3

19. Sokolova V, Faller J, Thai T, Engel L. Exploring relative health spillover effects within the context of dementia using the EQ-HWB-9. *Health Qual Life Outcomes*. Published online January 31, 2026. doi:10.1186/s12955-026-02486-z
20. Bailey C, Peasgood T, Michalowsky B, Engel L. The Psychometric Performance of the EQ Health and Wellbeing in Caregivers of Persons Living With Dementia. *Value in Health*. 2025;28(8). doi:10.1016/j.jval.2025.04.2168
21. Purba FD, Putri GHF, Mudita PGPR. Psychometric Performance of EQ-HWB and EQ-HWB-9 Self-Complete and Interviewer-Administered Versions in Literate, Low-Literacy, and Patient Populations in Indonesia. *Value Health Reg Issues*. 2026;53. doi:10.1016/j.vhri.2025.101543
22. Xu RH, Yang C, Rencz F. Exploring recall periods for EQ-5D-5L and EQ-HWB-9: a Hong Kong population study. *Value in Health*. Published online January 2026. doi:10.1016/j.jval.2026.01.001
23. Nasser SC, Pickard AS, Nazari JL, Kuharic M. Frequency and severity response scales for pain and discomfort: psychometric insights from EQ-HWB. *Quality of Life Research*. 2025;34(8). doi:10.1007/s11136-025-04003-z
24. Kinchin I, Engel L, Rencz F. A Comparative Study of Health and Well-Being Measures in Ireland Using EQ Health and Wellbeing (EQ-HWB) and its Short Version, EQ-5D-5L, and ICEpop Capability Measure for Adults (ICECAP-A). *Value in Health*. 2025;28(8):1268-1279. doi:10.1016/j.jval.2025.04.2160
25. Xu RH, Yang C, Rencz F. A quantitative exploration of the performance of a potential 'health and wellbeing' visual analogue scale (EQ-HWB VAS) for the EQ-HWB-9. *Value in Health*. Published online March 2026. doi:10.1016/j.jval.2026.02.013
26. Metyvinyi Z, Plázár D, Becker K, Medvecz M, Rencz F. Qualitative testing of the EQ-HWB-9 with a 'blank' EQ-HWB VAS and EQ-5D-5L with skin bolt-ons in a rare genetic skin disease population. *Quality of Life Research*. 2026;35(3). doi:10.1007/s11136-026-04177-0
27. Kuharic M, Mulhern B, Sharp LK, Turpin RS, Pickard AS. Comparison of the EQ-HWB and EQ-HWB-S With Other Preference-Based Measures Among United States Informal Caregivers. *Value in Health*. 2024;27(7):967-977. doi:10.1016/j.jval.2024.03.003
28. Mookink LB, Elsmann EBM, Terwee CB. COSMIN guideline for systematic reviews of patient-reported outcome measures version 2.0. *Quality of Life Research*. 2024;33(11). doi:10.1007/s11136-024-03761-6
29. Peasgood T, Mukuria C, Brazier J, et al. Developing a New Generic Health and Wellbeing Measure: Psychometric Survey Results for the EQ-HWB. *Value in Health*. 2022;25(4). doi:10.1016/j.jval.2021.11.1361
30. Nasreddine ZS, Phillips NA, Bédirian V, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *J Am Geriatr Soc*. 2005;53(4):695-699. doi:10.1111/j.1532-5415.2005.53221.x
31. Herdman M, Gudex C, Lloyd A, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Quality of Life Research*. 2011;20(10):1727-1736. doi:10.1007/s11136-011-9903-x
32. Xie S, Wang D, Wu J, Liu C, Jiang W. Comparison of the measurement properties of SF-6Dv2 and EQ-5D-5L in a Chinese population health survey. *Health Qual Life Outcomes*. 2022;20(1). doi:10.1186/s12955-022-02003-y

33. Weng G, Hong Y, Luo N, et al. Comparing EQ-5D-3L and EQ-5D-5L in measuring the HRQoL burden of 4 health conditions in China. *The European Journal of Health Economics*. 2023;24(2):197-207. doi:10.1007/s10198-022-01465-7
34. Pan CW, Zhang RY, Luo N, et al. How the EQ-5D utilities are derived matters in Chinese diabetes patients: a comparison based on different EQ-5D scoring functions for China. *Quality of Life Research*. 2020;29(11). doi:10.1007/s11136-020-02551-0
35. Ping W, Zheng J, Niu X, et al. Evaluation of health-related quality of life using EQ-5D in China during the COVID-19 pandemic. *PLoS One*. 2020;15(6):e0234850–e0234850.
36. Wong ELY, Xu RH, Cheung AWL. Measurement of health-related quality of life in patients with diabetes mellitus using EQ-5D-5L in Hong Kong, China. *Quality of life research*. 2020;29(7):1913-1921.
37. Wong ELY, Xu RH, Cheung AWL. Health-related quality of life in elderly people with hypertension and the estimation of minimally important difference using EQ-5D-5L in Hong Kong SAR, China. *The European journal of health economics*. 2020;21(6):869-879.
38. Wang W, Bian Q, Zhao Y, et al. Reliability and validity of the Chinese version of the Patient Health Questionnaire (PHQ-9) in the general population. *Gen Hosp Psychiatry*. 2014;36(5). doi:10.1016/j.genhosppsy.2014.05.021
39. Kroenke K, Spitzer RL, Williams JBW. The PHQ-9: Validity of a brief depression severity measure. *Journal of general internal medicine : JGIM*. 2001;16(9):606-613.
40. Sun J, Liang K, Chi X, Chen S. Psychometric properties of the generalized anxiety disorder scale-7 item (GAD-7) in a large sample of chinese adolescents. *Healthcare (Switzerland)*. 2021;9(12). doi:10.3390/healthcare9121709
41. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A Brief Measure for Assessing Generalized Anxiety Disorder: The GAD-7. *Archives of internal medicine (1960)*. 2006;166(10):1092-1097.
42. Xu RH, Keetharuth AD, Wang LL, Cheung AW ling, Wong EL yi. Psychometric Evaluation of the Chinese Recovering Quality of Life (ReQoL) Outcome Measure and Assessment of Health-Related Quality of Life During the COVID-19 Pandemic. *Front Psychol*. 2021;12. doi:10.3389/fpsyg.2021.663035
43. Fung SF. Psychometric evaluation of the Warwick-Edinburgh Mental Well-being Scale (WEMWBS) with Chinese University Students. *Health Qual Life Outcomes*. 2019;17(1). doi:10.1186/s12955-019-1113-1
44. Sun Y, Luk TT, Wang MP, et al. The reliability and validity of the Chinese Short Warwick-Edinburgh Mental Well-being Scale in the general population of Hong Kong. *Quality of Life Research*. 2019;28(10). doi:10.1007/s11136-019-02218-5
45. McDool E, Mukuria C, Peasgood T. Psychometric Performance of the EQ Health and Wellbeing Short in a United Kingdom Population Sample. *Value in Health*. 2024;27(9):1215-1224. doi:10.1016/j.jval.2024.05.007
46. Terwee CB, Bot SDM, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol*. 2007;60(1):34-42. doi:10.1016/j.jclinepi.2006.03.012
47. Kamper SJ, Maher CG, Mackay G. Global rating of change scales: A review of strengths and weaknesses and considerations for design. *Journal of Manual and Manipulative Therapy*. 2009;17(3). doi:10.1179/jmt.2009.17.3.163

48. Gwet KL. *Handbook of Inter-Rater Reliability: Volume 1: Analysis of Categorical Ratings (Handbook of Inter-Rater Reliability - The Definitive Guide to Measuring the Extent of Agreement Among Raters)*. Vol 1. 2021.
49. Nunnally JC. *Psychometric Theory*. 2d ed. New York : McGraw-Hill; 1978.
50. Janssen MF, Birnie E, Bonsel GJ. Evaluating the discriminatory power of EQ-5D, HUI2 and HUI3 in a US general population survey using Shannon's indices. *Qual Life Res*. 2007;16(5). doi:10.1007/s11136-006-9160-6
51. RL G, Gorsuch RL. *Factor Analysis*. 2nd ed. Hillsdale, N.J. : L. Erlbaum Associates; 1983.
52. P.Stevens J. *Applied Multivariate Statistics for the Social Sciences (5th Ed)*. Vol 77. 2009.
53. DeVellis RF. *Scale Development : Theory and Applications*. Fourth Edi. Los Angeles : SAGE; 2017.
54. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed. Hillsdale, N.J. : L. Erlbaum Associates; 1988.
55. Monteiro AL, Kuharic M, Pickard AS. A Comparison of a Preliminary Version of the EQ-HWB Short and the 5-Level Version EQ-5D. *Value in Health*. 2022;25(4). doi:10.1016/j.jval.2022.01.003
56. Kuharić M, Pickard AS, Mukuria C, Finch AP. The Measurement Properties of the EQ-HWB and the EQ-HWB-S in Italian Population: A Comparative Study With EQ-5D-5L. *Value in Health*. 2024;27(7). doi:10.1016/j.jval.2024.03.002

Table 1 EQ-HWB items ceiling, floor and test-retest reliability

EQ-HWB		Floor, n (%)			Ceiling, n (%)			Gwet's AC2/ICC	
modified	experimental	modified	experimental	p-value	modified	experimental	p-value	modified	experimental
1.Day-to-day activities *	Day-to-day activities *	11 (1.2)	8 (0.9)	0.628	506 (56.2)	524 (58.4)	0.122	0.90	0.92
2.Personal care	Personal care	4 (0.4)	4 (0.4)	1.000	693 (77.1)	701 (78.1)	0.395	0.97	0.92
3.Seeing	Seeing	2 (0.2)	3 (0.3)	1.000	550 (61.1)	515 (57.3)	0.009	0.92	0.95
4.Hearing	Hearing	5 (0.6)	3 (0.3)	0.683	673 (74.8)	655 (73.0)	0.184	0.97	0.96
5.Getting around inside or outside *	Getting around inside and outside *	3 (0.3)	4 (0.4)	1.000	692 (76.9)	687 (76.6)	0.931	0.97	0.95
6.Difficulty remembering	Trouble remembering	71 (7.9)	51 (5.7)	0.008	125 (13.9)	187 (20.8)	<0.001	0.73	0.83
7.Difficulty concentrating/thinking clearly *	Trouble concentrating/thinking clearly *	106 (11.8)	67 (7.5)	<0.001	148 (16.4)	205 (22.9)	<0.001	0.72	0.76
8.Hopeless	Nothing to look forward to	116 (12.9)	102 (11.4)	0.180	259 (28.8)	289 (32.3)	0.006	0.79	0.76
9.Lacked control over daily life *	No control over daily life *	85 (9.4)	88 (9.8)	0.828	270 (30.0)	310 (34.6)	<0.001	0.81	0.84
10.Excluded by others	Accepted by others	65 (7.2)	122 (13.6)	<0.001	353 (39.2)	169 (18.8)	<0.001	0.92	0.28
11.Negative feelings about self	Feel good about self	112 (12.4)	145 (16.2)	0.018	199 (22.1)	135 (15.1)	<0.001	0.79	0.46
12. Difficulty with enjoyable activities	Do things one wanted to do	19 (2.1)	164 (18.3)	<0.001	484 (53.8)	116 (12.9)	<0.001	0.93	0.47
13.Sleep		160 (17.8)			81 (9.0)			0.67	
14.Exhaustion*		149 (16.6)			118 (13.1)			0.67	
15.Loneliness*		121 (13.4)			214 (23.8)			0.80	
16.Unsupported		84 (9.3)			259 (28.8)			0.79	
17.Anxiety*		175 (19.4)			85 (9.4)			0.67	
18.Unsafe		66 (7.3)			270 (30.0)			0.82	

19.Frustration	107 (11.9)	183 (20.3)	0.75
20.Sadness/depression*	129 (14.3)	164 (18.2)	0.73
21.Coping	41 (4.6)	308 (34.2)	0.88
22.Pain (frequency)	29 (3.2)	188 (20.9)	0.73
23.Pain (severity)*	3 (0.3)	233 (25.9)	0.90
24.Discomfort (frequency)	32 (3.6)	183 (20.4)	0.74
25.Discomfort (severity)	6 (0.7)	188 (20.9)	0.87
EQ-HWB	0	0.8	0.77
EQ-HWB-9	0	2.4	0.73

Note: Items 1–12 were revised in the modified version compared to the experimental version.

* Items in the EQ-HWB-9.

P-value was estimated by the McNemar's test.

Item order does not follow the sequence in the EQ-HWB, as the table first presents the modified items and then the remaining ones.

Table 2 Results of the exploratory cross-instrument dimensionality (PCA)

	Mental health and cognition	Positive wellbeing	Physical health	Pain and discomfort	Social relationships
EQ-HWB (modified)					
Day-to-day activities *	0.41		0.61		
Difficulty with enjoyable activities			0.53		
Getting around inside or outside *			0.75		
Personal care			0.73		
Seeing			0.43		0.40
Hearing			0.60		
Sleep	0.47			0.45	
Exhaustion*	0.63				
Loneliness*	0.59				0.42
Unsupported	0.51				0.52
Difficulty remembering	0.43				
Difficulty concentrating/thinking clearly *	0.58				
Anxiety*	0.72				
Unsafe	0.49				
Frustration	0.70				
Sadness/depression*	0.70				
Hopeless	0.64	0.42			
Lacked control over daily life *	0.60				
Coping	0.55				
Excluded by others	0.45				0.55
Negative feelings about self	0.65				
Pain (frequency)				0.78	

Pain (severity)*		0.77
Discomfort (frequency)		0.62
Discomfort (severity)		0.70
EQ-5D-5L		
Mobility	0.67	
Self-care	0.68	
Usual Activities	0.62	
Pain/Discomfort		0.75
Anxiety/Depression	0.71	
GAD-7		
Feeling nervous, anxious, or on edge	0.80	
Not being able to stop or control worrying	0.80	
Worrying too much about different things	0.77	
Trouble relaxing	0.78	
Being so restless that it is hard to sit still	0.73	
Becoming easily annoyed or irritable	0.73	
Feeling afraid as if something awful might happen	0.63	
PHQ-9		
Little interest or pleasure in doing things	0.68	
Feeling down, depressed, or hopeless	0.74	
Trouble falling or staying asleep or sleeping too much	0.50	0.42
Feeling tired or having little energy	0.68	
Poor appetite or overeating	0.54	
Feeling bad about yourself or that you are a failure or have let yourself down	0.59	
Trouble concentrating on things, such as reading the newspaper or watching television	0.59	

Moving or speaking so slowly that other people could have noticed. Or the opposite	0.54		
Thoughts that you would be better off dead or of hurting yourself in some way	0.50		0.45
ReQoL-10			
I found it difficult to get started with everyday tasks	0.60		
I felt able to trust others		0.63	
I felt unable to cope	0.63		
I could do the things I wanted to do		0.72	
I felt happy		0.69	
I thought my life was not worth living	0.50		
I enjoyed what I did		0.76	
I felt hopeful about my future		0.76	
I felt lonely	0.57		
I felt confident in myself		0.78	
Physical health	0.51		0.55
SWEMWBS			
I've been feeling optimistic about the future		0.76	
I've been feeling useful		0.76	
I've been feeling relaxed	0.46	0.67	
I've been dealing with problems well		0.74	
I've been thinking clearly		0.73	
I've been feeling close to other people		0.73	
I've been able to make my own minds about things		0.73	

* Items in the EQ-HWB-9

Table 3 Results of convergent validity

EQ-HWB (items)	Other instruments or their items	Correlation coefficients	
		experimental	modified
Day-to-day activities*	EQ-5D-5L Usual activities	0.72	0.62
Difficulty with enjoyable activities	ReQoL I enjoyed what I did	0.36	-0.42
Getting around inside or outside*	EQ-5D-5L Mobility	0.59	0.55
Personal care	EQ-5D-5L Self-care	0.63	0.52
Difficulty concentrating/thinking clearly*	PHQ-9 Trouble concentrating on things	0.71	0.68
Hopeless	PHQ-9 Feeling down, depressed, or hopeless	0.71	0.71
Lacked control over daily life*	ReQoL-10 I felt unable to cope	0.66	0.59
Excluded by others	SWEMWBS I've been feeling close to other people	0.30	-0.49
Negative feelings about self	SWEMWBS I've been feeling useful	0.32	-0.57
Sleep	PHQ-9 Trouble falling or staying asleep or sleeping too much		0.76
Exhaustion*	PHQ-9 Feeling tired or having little energy		0.71
Loneliness*	ReQoL-10 I felt lonely		0.78
Unsupported	ReQoL-10 I felt able to trust others		-0.40
Anxiety*	GAD-7 Feeling nervous, anxious, or on edge		0.70
Unsafe	GAD-7 Feeling afraid as if something awful might happen		0.61
Frustration	PHQ-9 Feeling bad about yourself		0.67
Sad/depressed*	PHQ-9 Feeling down, depressed, or hopeless		0.75
Coping	ReQoL-10 I felt unable to cope		0.59
Pain (frequency)	EQ-5D-5L Pain/Discomfort		0.65
Pain (severity)*	EQ-5D-5L Pain/Discomfort		0.68
Discomfort (frequency)	EQ-5D-5L Pain/Discomfort		0.59

Discomfort (severity)	EQ-5D-5L Pain/Discomfort	0.67
EQ-HWB LSS	EQ-5D-5L LSS	-0.80
	EQ VAS	-0.66
	GAD-7	0.83
	PHQ-9	0.90
	ReQoL-10 LSS	0.78
EQ-HWB-9 LSS	SWEMWBS LSS	-0.67
	EQ-5D-5L LSS	-0.77
	EQ VAS	-0.65
	GAD-7	0.82
	PHQ-9	0.89
	ReQoL-10 LSS	0.77
	SWEMWBS LSS	-0.66

* Items in the EQ-HWB-9;

All correlation coefficients are statistically significant.

Pearson's correlation coefficients were calculated for scale-level correlation analysis; and Spearman's correlation coefficients were calculated for item-level correlation analysis.

Table 4 Results of known-group analysis on scale-level

	N	EQ-HWB LSS (mean, SD)	EQ-HWB-9 LSS (mean, SD)
Anxiety/depression duration, years			
<3	377	58.5 (21.3)	56.4 (22.3)
3-9	372	60.9 (21.0)	60.1 (21.6)
>9	86	66.3 (23.7)	66.1 (24.3)
p-value		0.009	<0.001
η^2		0.01	0.02
General health status[#]			
Good	344	76.5 (15.4)	75.8 (16.5)
Fair	424	52.5 (17.2)	50.4 (17.8)
Poor	130	42.0 (18.2)	40.8 (18.3)
p-value		<0.001	<0.001
η^2		0.39	0.39
Co-morbidity with non-mental health chronic conditions			
Yes	232	59.5 (20.9)	50.1 (22.1)
No	663	64.4 (23.1)	66.0 (23.2)
p-value		0.003	<0.001
Cohen's d		0.23	0.30
Co-morbidity with mental health conditions			
Yes	279	57.5 (22.8)	54.5 (24.1)
No	283	71.3 (19.6)	70.3 (20.9)
p-value		<0.001	<0.001
Cohen's d		0.65	0.70
Psychiatric medication prescribed			
Yes	617	56.1 (20.8)	53.3 (21.7)
No	277	62.9 (21.6)	62.1 (22.3)
p-value		<0.001	<0.001
Cohen's d		0.32	0.40
GAD-7 severity groups			
Minimal	479	74.6 (14.4)	73.47 (15.87)
Mild	251	47.8 (13.3)	46.45 (13.29)
Moderate	114	38.2 (15.2)	35.86 (15.50)
Severe	40	30.4 (14.1)	27.95 (14.41)
p-value		<0.001	<0.001
η^2		0.56	0.55
PHQ-9 severity groups			
Minimal or none	228	85.4 (8.6)	85.4 (9.8)
Mild	188	68.6 (10.0)	66.8 (10.9)
Moderate	195	54.6 (10.9)	52.3 (11.4)
Moderately severe	181	42.3 (10.1)	40.7 (10.4)
Severe	104	30.2 (12.5)	28.3 (13.0)
p-value		<0.001	<0.001
η^2		0.77	0.76

Note: The values do not add up to 900 due to missing values.

[#] Reported based on their responses to the 1-item global health question.

Figure 1 Response distribution of EQ-HWB; * Indicates items reworded from positive to negative framing.

One missing response was recorded for the item “personal care,” “exhaustion,” and “discomfort (frequency).”

Proportions less than 2% were not shown in the figure.

Figure 2 Known-group validity assessment on 12 modified EQ-HWB items.

Journal Pre-proof

