White Rose Research Online

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk
Universities of Leeds, Sheffield and York

Deposited via The University of Sheffield.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/id/eprint/243021/

Version: Published Version

Article:
Bailey, C., Trapani, K., Davies, J.N. et al. (2026) Validation of the EQ-HWB-9 in a mental health sample and an investigation of modifications to items. Health and Quality of Life Outcomes, 24 (120). ISSN: 1477-7525

https://doi.org/10.1186/s12955-026-02579-9

Reuse
This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) licence. This licence only allows you to download this work and share it with others as long as you credit the authors, but you can't change the article in any way or use it commercially. More information and the full terms of the licence here: https://creativecommons.org/licenses/

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

UNIVERSITY OF LEEDS
University of Sheffield
UNIVERSITY of York

Health and Quality of Life Outcomes

RESEARCH

Open Access

Validation of the EQ-HWB-9 in a mental health sample and an investigation of modifications to items

Cate Bailey[1,2*], Karen Trapani[3], Jonathan N. Davies[2], Nicholas T. Van Dam[2], Julieta Galante[2] and Tessa Peasgood[4]

Abstract

Background The EQ Health and Wellbeing (EQ-HWB) is a new, generic instrument designed to evaluate quality-of-life across health, public health, and social care settings. The short form comprises nine items (EQ-HWB-9) and validation across diverse populations and contexts is required. We aimed to investigate the validity of the EQ-HWB-9 in an international sample of adults experiencing poor mental health who downloaded a meditation app. We further examined the impact of four item-level modifications on psychometric performance, including investigating a potential ordering effect for the 'activities' item (hypothesised in prior studies) and three minor changes to response options.

Methods The current study was embedded in a larger trial examining engagement with meditation via a free, downloadable app. Participants were randomised to complete the original (2022) and modified (2024) experimental version of the EQ-HWB-9. Psychometric evaluation included analyses of item distribution, known group and convergent validity, and responsiveness to change.

Results There were no differences in demographic characteristics between the EQ-HWB original and modified versions at baseline ($n=865$) or follow-up ($n=130$). All psychometric tests supported the validity of the EQ-HWB-9 in this population. We found an ordering effect for the activities item, where the activities item showed a greater level of difficulty and a wider distribution over response options when asked before the mobility item, rather than after. There were no observable differences between versions for the other modifications.

Conclusions These findings add to growing literature supporting the EQ-HWB-9 as a suitable instrument for measuring quality-of-life across a range of settings. When the 'activities' item was presented first, as in the modified version, participants appeared to interpret the item more broadly, in line with the developers' intentions. Accordingly, our results support changing the item order of the first two items. Other modifications had little impact on outcomes, suggesting that further qualitative research will be required to inform decisions about their inclusion. Results from this study provide support towards finalisation of the instrument. The development of country-specific value-sets is now critical to support its application in economic evaluation.

*Correspondence:
Cate Bailey
cate.bailey@unimelb.edu.au

Full list of author information is available at the end of the article

© The Author(s) 2026. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

Keywords EQ-HWB, EuroQol health and wellbeing, Quality-of-life, Instrument validation, Psychometrics, Mental health

Introduction

The EQ Health and Wellbeing (EQ-HWB) is a recently developed generic instrument designed to evaluate quality-of-life across health, public health, and social care settings [1]. Two versions exist: a 25-item full instrument (EQ-HWB) and a 9-item short form (EQ-HWB-9). Recent studies have found that the EQ-HWB-9 has robust psychometric outcomes in caregiver populations [2–7] and in general populations in China, UK and Australia [7–10]. Psychometric testing has included tests of known-group and convergent validity [2, 3, 6, 7, 9] and responsiveness to change [2, 11]. A pilot UK value set for the EQ-HWB-9 ('experimental', 2022, v1.1) has been published [12], enabling calculation of quality-adjusted life years (QALYs). Further testing of the instruments in a range of populations is required, particularly in those populations for which the instrument was intended, such as caregivers, social care users, and in mental health [13].

Mental health disorders are a leading cause of global disease burden, affecting individuals' functioning, relationships, and quality-of-life across the life course [14]. Increasing rates of depression, anxiety, and stress-related conditions, combined with limited access to timely and effective mental healthcare, have led to growing interest in digital health interventions [15], with recent research finding that mindfulness apps have small but significant effects in reducing anxiety and depression [16]. Prior research shows that meditation app users tend to have poorer mental health status than the general population; in a recent study of 886 meditators in the United States, nearly one-quarter (24%) met criteria for a moderate or severe mental disorder [17].

Understanding the impact of meditation apps is important not only for clinical practice but also for informing decisions about resource allocation. Economic evaluation of these interventions requires robust, preference-based measures of health and wellbeing that can capture relevant domains of impact [18]. Digital mental health interventions often target psychosocial outcomes that align with EQ-HWB-9 domains, particularly anxiety, low mood, fatigue, and reduced sense of control. Therefore, evaluating the psychometric performance of the EQ-HWB-9 in users of meditation apps can inform its suitability for economic evaluations of digital mental health programs.

The EQ-HWB instruments are in the final stages of development before being released for general use. Finalisation of the short-form is a priority for the developers due to its relevance for economic evaluation. A range of modifications are currently being considered such as changes to item ordering, and minor wording changes to items and response options. In our previous qualitative work, we observed context effects, in that that some participants perceived the activities item as a physical item when presented after mobility; however, when queried, participants stated much of their difficulty with day-to-day items was caused by psycho-social reasons such as exhaustion and caring responsibilities [2]. When participants were asked about the order of items, most participants supported changing the item order to present activities first [4]. We were therefore interested in whether we could detect whether participants perceived the activities item more broadly if it was presented prior to the mobility item using quantitative analysis. Testing the proposed modifications to the EQ-HWB-9 is a current priority for instrument developers.

The aim of this study was to use psychometric testing to assess whether the EQ-HWB-9 was a suitable quality-of-life measure in a sample of adults who downloaded a mindfulness meditation app. We examined response distributions, mean differences and correlations between items, known-group validity at the sum-score and index-score level, convergent validity with measures of mental health symptoms and wellbeing, and responsiveness to change in this population. Because of the large sample size expected for the study, we were able to test a set of modifications being considered for the EQ-HWB-9 by randomising participants to the original or modified version and investigate any differences in outcomes whilst maintaining sufficient power for the psychometric analyses. Hence, we further aimed to compare the EQ-HWB-9 ('experimental', 2022, v1.1) version to the EQ-HWB-9 ('experimental modified', 2024, v1.2) version of the EQ-HWB-9 to explore the effect of the modifications on the instruments' psychometric properties in this sample.

Method

Study design and procedure

On downloading the meditation app, an in-app notification invited users to participate in a research study. This notification then directed participants outside the app to information on the study, a consent form, and a link to the online survey. Data were collected between July and September 2024. Participants completed brief self-report measures in English on day 0 (baseline) and day 28 (post-intervention). Participants needed to be over 18 years of age. The current study was a "piggy-back" study; the aim of the main study was to investigate participant engagement with a meditation app [33].

Materials

Data collection included a range of demographic variables (age, gender, educational attainment, family income, employment status and geographic location) and information on mental health and app use (past meditation practice, the use of mindfulness apps, unpleasant experiences from meditation practice, the functional impact of reported unpleasant experiences, ever or currently receiving professional mental health support). Participants provided information about their combined total family income by selecting an income band from a 5-tier country specific income range. The income range that the participant was presented with was based on their country of residence. We coded this into a common 5-tier income range where the lowest band was mapped to the value "Lowest" and the highest band was mapped to the value "Highest" and the intermediate bands were labelled "2", "3" and "4" to allow inclusion of this variable in the analyses.

Data from five instruments was collected. The EQ-HWB-9 included 9-items covering physical (activities, mobility, pain) and psychosocial items (exhaustion, loneliness, cognition, anxiety, sadness/depression, lack of control) over the last seven days. Each item has five response options. EQ-HWB-9 sum scores were calculated by summing individual item scores. The index-score was calculated by applying a pilot EQ-HWB-9 value-set from the UK [12]. The Kessler Psychological Distress Scale (K10) [19] is a 10-item questionnaire of psychological distress measured over the previous 30-days. The scale ranges from 1 (none of the time) to 5 (all of the time). K10 sum scores were calculated by summing all items. Cut points were defined as likely to be well or have a mild disorder (< = 24) or likely to have a moderate or severe disorder (> 25) [20]. The Short Warwick-Edinburgh Mental Wellbeing Scale (SWEMWBS) [21] assesses positive facets of mental health on a 7-item, 5-point scale: 1 (None of the time), 2 (Rarely), 3 (Some of the time), 4 (Often) and 5 (Always), and is calculated by summing all items. Data were adjusted to produce a metric score [22]. Cut-points for three groups were defined as low (7- 19.3), medium (20–27.0) and high (28.1–35.0) wellbeing [23]. The Perceived Stress Scale (PSS-4) [24] was used to assess the participant's perceived stress in the past month with four-items across five response options 0 (Never), 1(Almost never), 2 (Sometimes), 3 (Fairly often), 4 (Very often). Items 2 and 3 were reverse-scored and the items were then summed. Cut-points were defined as no or low stress (< 6) compared to higher stress (> = 6) [25].The Satisfaction with Life Survey (1-item) (SWLS1) [26] is an abbreviated measure of the Satisfaction with Life Survey with a single-item asking participants to rate life satisfaction on a 7-item scale ranging from 1 (extremely dissatisfied) to 7 (extremely satisfied).

Self-reported utilisation of professional mental health support was measured through the following question: *"Have you ever received professional mental health support? This could include support from a primary care provider, GP, family medicine doctor, therapist, social worker, psychologist or psychiatrist, online or in person" (No / Yes – received support in the past / Yes – am receiving support now).* "Current mental health support" was coded as participants who were only currently receiving mental health support versus all others, and "Any mental health support" for participants who reported current or previous use versus 'no'.

Sleep was measured using the EQ-HWB long form Item 9, with permission from the developers, *"In the last 7 days, did you have problems with your sleep?"* on a five-point response scale (noting that sleep is not included in the short-form). "Sleep problems" were recoded to yes (sometimes/often/most of the time) and no (none of the time/only occasionally). Unpleasant experiences with meditation were measured using the question: "*Have you ever had any particularly unpleasant experiences (e.g., anxiety, fear, distorted emotions or thoughts, altered sense of self or the world) during your overall meditation practice?*") (yes/no). Participants who responded affirmative were asked to respond to an additional statement "*My meditation-related challenging, difficult or distressing experiences impaired my ability to function*" by selecting from a 4-item answer scale (not at all / somewhat / moderately / severely). We coded a binary variable "*Functional impairment from meditation*" for participants who responded with somewhat, moderately or severely, compared to those who answered not at all and who had indicated no unpleasant experiences in the previous question.

Missing data

The study dataset included baseline survey responses from 1021 participants. We excluded survey responses if participants had not completed the five instruments (156 records) for a final baseline sample of 865. We had 130 participants with EQ-HWB-9 data at follow-up. Missing data analysis suggested that participants who did not complete data collection were more likely to be male, and significantly but with small percentage differences for education (slightly lower) and occupation (slightly less likely to work full time), and slightly younger (not significant) (Table S8).

Differences between original and modified versions of the EQ-HWB-9

Participants were randomised to complete either the EQ-HWB-9 ('experimental', 2022, v1.1) or the EQ-HWB-9 ('experimental modified', 2024, v1.2) version of the instrument. We tested the following modifications: (1) swapping the order of the first two items (in the modified

version activity now preceded mobility); (2) adjusting the third response option of the difficulty response scale (items 1 and 2) from "*Some* difficulty" to "*Moderate* difficulty"; (3) Adjusting the second response option of the frequency response scale (items 3–8) from "*Only occasionally*" to "*A little of the time*"; and (4) adjusting the wording of Item 8 from "In the last 7 days, did you feel you *had no* control over your day-to-day life? (e.g. had no choice *to do things or have things done for you as you liked and when you wanted*)" to "In the last 7 days, did you feel you *lacked* control over your day-to-day life? (e.g. had no choice *over what you did or how things were done for you*)".

Statistical analysis

Analyses were conducted separately for groups completing the original experimental and the modified experimental versions, except for the responsiveness to change anlayses where the sample size was too small. Demographic characteristics were compared between participants in the two version groups using chi-squared and *t*-tests. Response distributions were calculated for item scores using frequency and percentages at baseline and follow-up. We compared mean scores of the EQ-HWB-9 items across the two versions using *t*-tests. We correlated EQ-HWB-9 items to investigate differences between versions. We used the CORTESTI module in STATA to test the equality of item-level correlation coefficients between original and modified versions [27].

Known group validity analysis was conducted at the sum-score and index-score levels using *t*-tests (binary outcomes)and one-way ANOVAs (3 or 4 groups). EQ-HWB-9 level sum-scores and preference-weighted scores were treated as continuous measures. Effect sizes were measured using Cohen's *d* (0.2 small, > 0.5 moderate, < 0.8 large [28]). We hypothesised that lower quality-of-life (higher EQ-HWB-9 sum-scores, lower EQ-HWB-9 index-scores) would be associated with poorer mental health status (K10) [2, 8], lower well-being (SWEMWBS) [11, 29], higher psychological stress (PSS-4) and lower satisfaction with life (SWLS1), and sleep problems [8]. We hypothesised that participants receiving current or any professional mental health support would have lower quality-of-life than participants not receiving professional mental health support, on the basis that they were likely to have had mental health issues in order to seek support (acknowledging that this hypothesis could be reversed if mental health support was highly successful, and those not receiving support had needed it). We hypothesised that unpleasant experiences and functional impairment associated with meditation practice would be associated with lower quality-of-life, although committed meditators may be more likely to have experienced and unpleasant effect.

We used Spearman's rank-order correlation (coefficient) for ordinal data to compare the EQ-HWB-9 and K10, SWEMWBS and PSS4 items and total scores. We defined correlation strength as 0–0.09 no correlation, 0.1–0.29 weak, 0.3–0.49 moderate, and ≥ 0.5 strong [28]. Hypotheses were determined a priori with consensus (TP and CB) for items theoretically likely to be correlated.

For the responsiveness to change analyses we calculated the EQ-HWB-9 change in sum-score between baseline and follow-up (follow-up minus baseline) for the 130 cases for which we had follow-up data. Change scores were calculated for the K10, SWEMWBS and PSS4 instruments in the same way, and then recoded into three groups for 'improved', 'no change', and 'worsened' scores based on the distribution of the scores, as shown in Table S1. For the EQ-HWB-9 to be seen as responsive to change, we would expect that deteriorating scores on the K10, SWEMBWS, PSS4 and SWLS1 would be associated with deteriorating quality-of-life scores and vice versa. Unfortunately, the versions were inadvertently "re-randomised" by the data collection team at follow-up. The re-randomisation created four groupings, such that we had half the number of participants we could track across time with the same randomisation: ie (1) experimental version at baseline and follow-up, (2) experimental version at baseline and modified version at follow-up, (3) modified version at baseline and experimental version at follow-up, (4) modified version at baseline and follow-up. In the manuscript, we have presented all EQ-HWB-9 data combined due to the low sample-size, and presented the subgroup analysis (by the four groups) in supplementary files. We included correlations between EQ-HWB-9 change scores and change scores from the other outcome measures.

Results

Comparison of baseline differences between original and modified samples

The baseline sample of 865 participants was predominantly below 39 years of age (79%), 55% female, well educated (only 3% did not complete school), and employed full time (41%) (Tables S2a/b). As expected, there were high rates of psychological distress, with 45.2% of participants likely to have moderate to severe psychological disorders on the K10. Participants were from Europe (44%) (Table S2c for full list), United States (29%), Australia (9%), United Kingdom (8%) Canada (8%) and New Zealand (2%). There were no statistically significant differences in baseline demographic characteristics or mean differences in outcome measures (K10, SWEMWBS, PSS4 and SWLS1) between participants randomised to the original version of the EQ-HWB-9 ($n = 450$) compared to the modified version ($n = 415$) (Tables S2a/b).

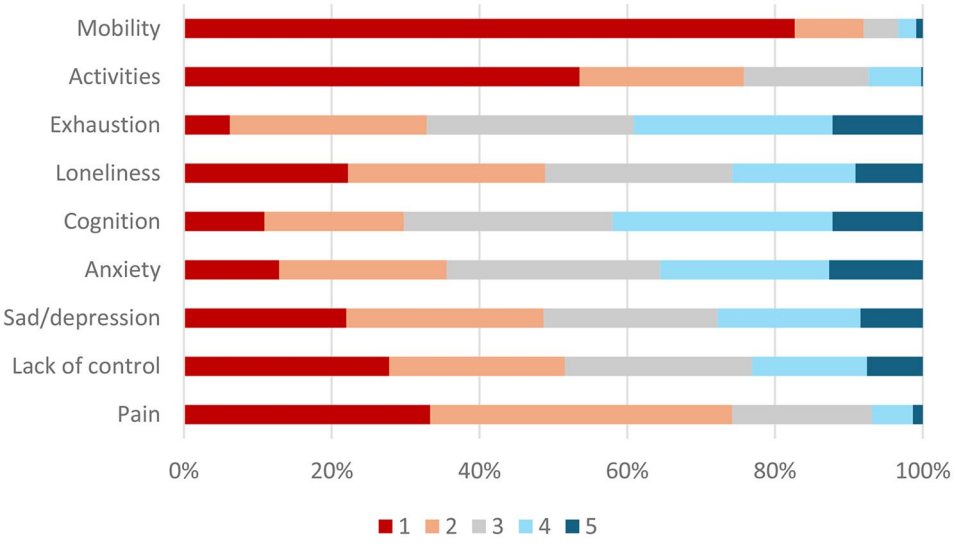


Fig. 1 Distribution of EQ-HWB-9 scores for the original version at baseline ($n = 450$)

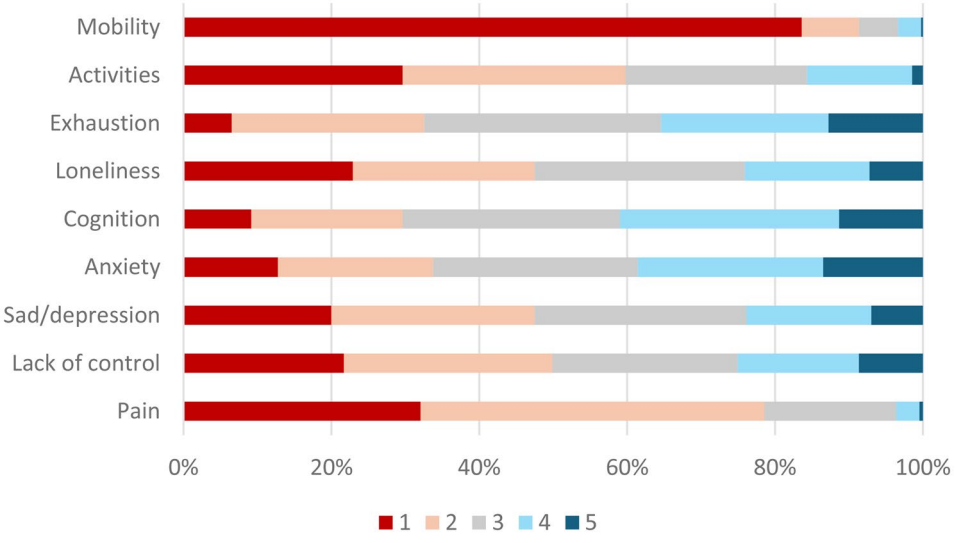


Fig. 2 Distribution of EQ-HWB-9 scores for the modified version at baseline (includes all modifications) ($n = 415$)

The follow-up response rate was low at 130 participants (15.0%).

EQ-HWB-9 item distributions for original and modified versions

Figures 1 and 2 show the distributions of the original and modified versions of the instrument. Numbers and percentages with percentage difference across the response options are displayed in Table S3.

Note x-axis numbers correspond to response options, which are different across items

Comparing items; mean scores comparisons and correlations

The mean scores of the original version of the activities item (1.79, SD = 0.98) were statistically significantly lower than the mean scores for the modified version (2.28, SD = 1.08) (mean difference = 0.50, t (df) = 7.06 (881), $p<.001$, Cohen's d = 0.48), suggesting an item ordering effect. There were no statistically significant differences between the original and modified versions for any other items (Table S4). Mean results are displayed in Fig. 3.

*$p<.001$.

We compared correlations between EQ-HWB-9 items for the original and modified versions and highlighted any correlation pairs, as shown in Table 1. This analysis suggests a pattern of differences between original and

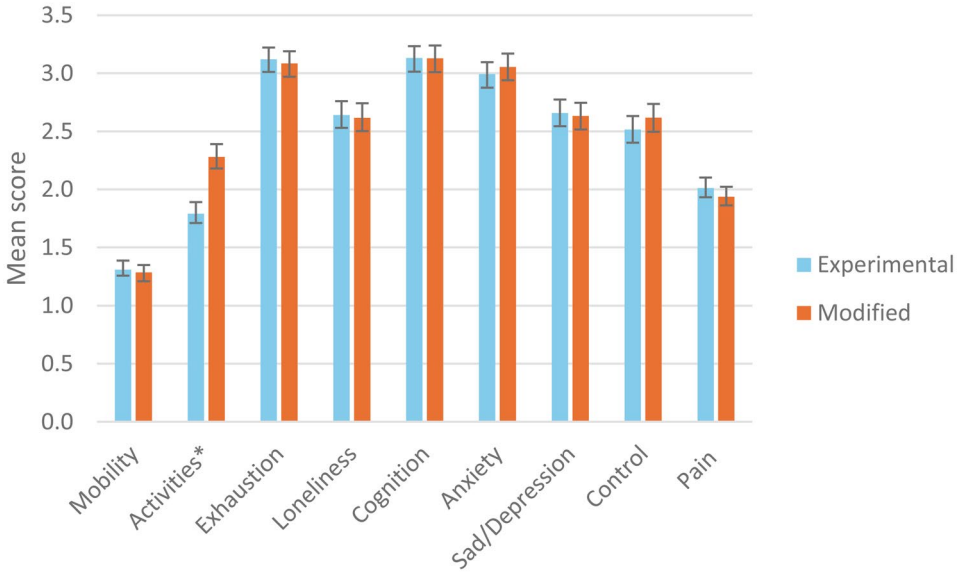


Fig. 3 Mean scores of EQ-HWB-9 scores for original and modified versions with 95% confidence intervals

modified versions of the 'activities' item with other items. For the modified version of the instrument, the 'activities' item had lower correlations with mobility and higher correlations with exhaustion, cognition and sad/depression items.

Known-group analyses

We conducted known-group analysis using *t*-tests across a range of variables, shown in full in Table S5a for the sum-score, and Table S5b for the index-scores. In the original and modified versions, all tests were significant in hypothesized directions, with *p*-values less than 0.001 for sum-score analyses, except current mental health support in the original version, which was significant at less than 0.01. Cohen's *d* scores are summarised in Table 2 to investigate differences between original and modified versions; results appear to be similar across versions. There were no apparent differences between the original and modified versions of the instrument in the known-group analyses by 3 and 4 groups, as shown in Table S5c.

Convergent validity analysis

All hypothesised correlations between EQ-HWB-9 and the K10 (Table S6a), the SWEMWBS (Table S6b), and the PSS4 (Table S6c) were above the threshold of a moderate correlation (over 0.3) for both the original and modified versions; hence we found no differences between instrument versions in this regards.

Responsiveness to change

Responsiveness to change analyses for the $n = 124–130$ participants with follow-up EQ-HWB-9 data against the K10, SWEMWBS, PSS4 and SWLS1 were conducted. Original and modified combined results are shown in Table 3. All tests conformed to hypothesised directions. Reduced psychological distress (K10), reduced wellbeing (SWEMBWS) and reduced stress (PSS4) were all associated with improved quality-of-life (EQ-HWB-9). Correlations between the EQ-HWB change sum-scores and the change scores for other outcome measures were 0.551 for the K10, – 0.467 for the SWEMWEBS, 0.452 for the PSS4, and – 0.407 for the SWLS1. Subgroup results for the original and modified versions are contained in in Tables S7a/b. Due to the small sample sizes across the groups, some tests lacked statistical significance, and comparisons of the effect sizes between the two versions of the instrument could not be conducted.

Discussion

This study examined the validity of the EQ-HWB-9 for use in an adult sample with higher mental health concerns than the general population who downloaded a meditation app. Our assumption that the population would have higher mental health problems was confirmed, with 45% of participants classified as likely to have a moderate to severe disorder compared to 14% in the Australian population in 2022 [30]. Overall, item distribution for both versions was similar to previous studies [2–5]: participants reported higher mean scores for the items *feeling exhausted*, *cognition* and *anxiety* and lower problems on the 'physical' items (*mobility*, *activities*, *pain*). Both versions discriminated well between groups on the known-groups validity analyses, with large observed between-group effect sizes (< 0.8) for psychological distress, low mental wellbeing, stress, satisfaction with life, sleep problems, and meditation-related functional impairment, and small-to-moderate between-group effect sizes for unpleasant experiences

Table 1 Correlations between EQ-HWB-9 items at baseline for the original and modified versions and the significance of the difference

		Mobility	Activities	Exhaustion	Loneliness	Cognition	Anxiety	Sad/ depression	Lack of control	Pain
Mobility	Original	1.000***								
	Modified	1.000***								
	p^ (diff)	n/a								
Activities	Original	**0.448*** **	1.000***							
	Modified	**0.305*** **	1.000***							
	p^ (diff)	**0.014 (0.143)**	n/a							
Exhaustion	Original	0.270***	**0.435*** **	1.000***						
	Modified	0.308***	**0.566*** **	1.000***						
	p^ (diff)	0.544 (-0.038)	**0.010 (-0.131)**	n/a						
Loneliness	Original	0.156***	0.397***	0.398***	1.000***					
	Modified	0.167***	0.416***	0.407***	1.000***					
	p^ (diff)	0.879 (-0.011)	0.739 (-0.019)	0.875 (-0.009)	n/a					
Cognition	Original	0.191***	**0.416*** **	0.492***	0.484***	1.000***				
	Modified	0.123*	**0.527*** **	0.512***	0.403***	1.000***				
	p^ (diff)	0.307 (0.068)	**0.036 (-0.111)**	0.695 (-0.020)	0.139 (0.081)	n/a				
Anxiety	Original	0.187***	0.329***	0.450***	0.503***	0.480***	1.000***			
	Modified	0.128**	0.416***	0.465***	0.412***	0.546***	1.000***			
	p^ (diff)	0.376 (0.059)	0.139 (-0.087)	0.781 (-0.015)	0.091 (0.091)	0.189 (0.077)	n/a			
Sad/depression	Original	0.193***	**0.400*** **	0.457***	0.635***	0.436***	0.660***	1.000***		
	Modified	0.233***	**0.515*** **	0.498***	0.631***	0.468***	0.579***	1.000***		
	p^ (diff)	0.540 (-0.040)	**0.033 (-0.115)**	0.437 (-0.041)	0.922 (0.004)	0.556 (-0.032)	0.054 (0.081)	n/a		
Lack of control	Original	0.195***	0.411***	0.406***	0.423***	0.485***	0.470***	0.476***	1.000***	
	Modified	0.184***	0.480***	0.470***	0.440***	0.512***	0.483***	0.491***	1.000***	
	p^ (diff)	0.867 (0.011)	0.207 (-0.069)	0.246 (-0.064)	0.760 (-0.017)	0.600 (0.027)	0.806 (-0.013)	0.774 (-0.015)	n/a	
Pain	Original	0.339***	0.305***	0.341***	0.211***	0.189***	0.185***	0.182***	0.236***	1.000***
	Modified	0.292***	0.356***	0.356***	0.165***	0.304***	0.241***	0.226***	0.189***	1.000***
	p^ (diff)	0.445 (0.047)	0.402 (-0.051)	0.803 (-0.015)	0.485 (0.056)	0.073 (-0.115)	0.390 (-0.056)	0.501 (-0.044)	0.471 (0.050)	n/a

^ *p*-value for comparison between correlation coefficients. *p*-values and correlations are bolded if the *p*-value was less than .05. Differences between original and modified correlations are in brackets beside the *p*-values. For the correletions, significance is measured as * = $p<.05$, ** = $p<.01$, *** = $p<.001$

Table 2 Comparisons of Cohen's *d* scores between the original and modified versions, for sum-score and index-scores

	Sum-scores (d)			Index-scores (d)		
Known groups	**Original**	**Modified**	**Difference**	**Original**	**Modified**	**Difference**
Psychological Distress (K10)	-2.00	-1.93	0.07	1.82	1.77	-0.05
Low mental wellbeing (SWEMWBS)	-1.44	-1.42	0.02	1.28	1.33	0.05
Psychological Stress (PSS4)	-1.43	-1.41	0.02	1.17	1.24	0.07
Satisfied with life (SWLS1)	1.02	1.06	0.04	-0.97	-0.99	-0.02
Current mental health support^	-0.32	-0.46	-0.14	0.29	0.49	0.20
Any mental health support^^	-0.43	-0.55	-0.12	0.38	0.54	0.16
Sleep problems	-0.96	-1.02	-0.06	0.87	1.00	0.13
Unpleasant experiences with meditation	-0.74	-0.48	0.26	0.71	0.51	-0.20
Functional impairment from meditation	-0.87	-0.67	0.20	0.85	0.74	-0.11

^current vs past or none; ^^current and past vs noneCohen's *d*; 0.2 = small, >0.5 = moderate, <0.8 = large(28)

Table 3 Responsiveness to change for the EQ-HWB sum-score (all data combined)

		#	Mean	SD	F(df)	*p*-value
K10					16.15 (2,124)	< 0.001
	Reduced mental distress	49	-2.59	4.12		
	No change	42	-0.79	3.76		
	Increased mental distress	36	2.61	4.72		
	Total	127	-0.52	4.66		
SWEMWBS					15.45 (2,122)	< 0.001
	Improved wellbeing	34	2.35	4.75		
	No change	43	-0.14	3.54		
	Reduced wellbeing	48	-2.88	4.41		
	Total	125	-0.51	4.7		
PSS4					11.39 (2,127)	< 0.001
	Reduced stress	36	-3.39	4.57		
	No change	55	0.42	3.75		
	Improved stress	39	0.87	4.74		
	Total	130	-0.50	4.63		
SWLS1					15.13 (2,121)	< 0.001
	Improved satisfaction	17	-4.00	3.61		
	No change	101	-0.39	4.24		
	Reduced satisfaction	6	7.00	6.07		
	Total	124	-0.52	4.72		

with mediation and mental health support. All convergent validity analyses conformed to a priori expectations. Sample sizes were low for the responsiveness to change analysis, but all results were in the expected directions. Thus, psychometric testing supports the use of the EQ-HWB-9 in this sample with higher mental health problems.

We examined whether there were differences in psychometric results between the original and modified versions of the EQ-HWB-9 instrument. No significant baseline or follow-up differences were observed between the randomised original and modified samples across any of the measured variables (age, gender, education, occupation, region, income), indicating that the randomised groups were comparable. Accordingly, any observed variation in the distributions or psychometric performance of the two versions would be expected to occur by chance, unless item modifications exerted a genuine influence on outcomes.

We found substantial evidence of an ordering effect for the first two EQ-HWB-9 items, by comparing the mean scores and by investigating the correlations. The modified activities item had significantly higher mean scores and a greater response spread than in the original version, suggesting that when the activities item is presented before the mobility item, participants may interpret 'activities' more broadly. This result concords with results from our prior qualitative work [2, 3], where participants stated that they were thinking of activities as being a physical item when the question came after 'mobility' as in the original version, rather than other important aspects that affected their ability to do their day-to-day activities, such as being too tired or depressed, or not being able to think well.

The changes in correlations between the modified activities item and exhaustion, cognition, sad/depression and mobility items compared to correlations in the original version support our hypotheses of a broader interpretation of activities when that item was presented first. There were significant differences in correlations between the original and modified versions of the activities item when compared to mobility, exhaustion, cognition, and sadness/depression, which concords with results from our previous qualitative studies, where participants referred to aspects of mental health, exhaustion and difficulty with thinking as explanations of why they found day-to-day activities difficult [2, 4]. This broader interpretation of the activities item was more common if it was asked before the mobility item; whereas, if asked after the mobility item, participants were more likely to interpret the item as referring to physical aspects.

There were no differences in psychometric outcomes for the three other modifications. It is possible that these changes were not sufficiently different from the original version as to affect psychometric outcomes in this study. Further quantitative and qualitative research will be needed to assess whether to incorporate theses changes in a revised version of the instrument. Other factors such as readability, accessibility, suitability for translations and other cultures may be considered.

Strengths and limitations

This study builds on previous work that has demonstrated the validity of the EQ-HWB-9 in a variety of settings. The large sample size enabled the randomisation of two versions of the EQ-HWB-9. The main limitation to the study was the follow-up sample size for the responsiveness to change analysis, and the study implementation error which meant that participants were randomised again at follow-up, rather than remaining in their original randomisation groups, which further lowered the sample size for analysis. This part of the analysis, therefore, can only be considered preliminary, with low statistical power. However, longitudinal data for the purposes of validating the EQ-HWB-9 has been limited, making this information highly valuable. Significant differences between correlations were set to less than 0.05, and it was not possible to determine whether these changes are practically meaningful or only statistically detectable. We note that mental health problems were self-reported, and that this was not a clinical sample. We used the UK Pilot value-set for both the EQ-HWB-9 ('experimental', 2022, v1.1) version and the EQ-HWB-9 ('experimental modified', 2024, v1.2) of the instrument and acknowledge that this value-set was developed for the original version and may not be as relevant for the modified version. We did not apply multiple testing correction given that group comparisons and responsiveness were specified a priori and our interest was in the magnitude of effect size rather than the binary significance threshold, thus limiting the need for correction. Correlation comparisons were exploratory and the large number of comparisons increases the risk of Type 1 errors. We acknowledge the limitations of using sum-scores, which gives equal weight to all items [31], whilst also noting that an unweighted approach may be suitable in some circumstances [32]. There are currently no value-sets for Australia so we used the UK pilot version [12]. Given the large number of countries included, participants may not have completed in the survey in their primary language; however, they downloaded the app in English so we can reasonably assume that they were fluent.

Conclusion

This study adds to the considerable literature evaluating the EQ-HWB-9 in a range of settings. In a large adult population sample of relatively young people, many with mental health concerns, we found that the EQ-HWB-9 had robust psychometric properties. We endorse switching the order of the first two items, as the results here quantify what we found in our previous qualitative studies. Other tested modifications had little effect on psychometric results. This study provides evidence to support the use of the instrument in this population as a generic instrument for measuring quality-of-life that the EQ-HWB Working Group can use to make final decisions on the EQ-HWB-9. The authors recommend that the instrument becomes available to the wider research community. The next critical step is to develop country-specific value-sets to enable the instrument's use in economic evaluation.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12955-026-02579-9.

Supplementary Material 1

Acknowledgements

We would like to acknowledge Julia Adams and Alex Burger for their help in data collection and management.

Author contributions

CB and TP contributed to the study conception and design. Material preparation and data collection was performed by JD, NVD, JG. Analysis were performed by CB and KT. The first draft of the manuscript was written by CB and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding

Funding of the data collection was supported by the Contemplative Studies Centre at the University of Melbourne, which was established by a philanthropic donation from the Three Springs Foundation Pty Ltd (Australian Company Number 625000562). The funders had no role in the study design or in its implementation, analysis, or reporting.

Data availability

The datasets used and/or analysed during the current study are available via OSF. [https://osf.io/3v897).

Declarations

Ethics approval

University of Melbourne Human Ethics Approval Committee - Reference 23969, in accordance with the Declaration of Helsinki. All participants consented to participate on entering the project.

Consent for publication

N/A. All data were de-identified prior to analysis.

Competing interests

The authors declare no competing interests.

Author details

[1]Melbourne Health Economics, School of Public Health, The University of Melbourne, 207 Bouverie St, Carlton, VIC 3035, Australia
[2]Contemplative Studies Centre, Melbourne School of Psychological Sciences, The University of Melbourne, Carlton, VIC, Australia
[3]Centre for Health Policy, School of Public Health, The University of Melbourne, Carlton, VIC, Australia
[4]Sheffield Centre for Health and Related Research, School of Medicine and Population Health, University of Sheffield, Sheffield, UK

Received: 27 January 2026 / Accepted: 23 June 2026
Published online: 30 June 2026

References

1. Brazier J, Peasgood T, Mukuria C, Marten O, Kreimeier S, Luo N, et al. The EQ-HWB: Overview of the Development of a Measure of Health and Wellbeing

and Key Results. Value Health. 2022;25(4):482–91. https://doi.org/10.1016/j.jval.2022.01.009.
2. Bailey C, Dalziel K, Constable L, Devlin NJ, Hiscock H, Skouteris H, 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. Eur J Health Econ. 2025;26(1):7–21. https://doi.org/10.1007/s10198-024-01688-w.
3. Bailey C, Dalziel K, Jones R, Hiscock H, Devlin NJ, Peasgood T, et al. The validity of the EuroQol Health and Wellbeing Short Version (EQ-HWB-S) instrument in parents of children with and without health conditions. PharmacoEconomics. 2024. https://doi.org/10.1007/s40273-024-01351-5.
4. Bailey C, Peasgood T. Content validity of the EQ-HWB in caregivers of children with health conditions. Patient. 2025. https://doi.org/10.1007/s40271-025-00749-3.
5. Bailey C, Peasgood T, Michalowsky B, Engel L. The psychometric performance of the EQ-HWB in caregivers of persons living with dementia. Value Health. 2025. https://doi.org/10.1016/j.jval.2025.04.2168.
6. Kuharic M, Mulhern B, Sharp LK, Turpin RS, Pickard AS. Comparison of EQ Health and Well-Being Long and Short with other preference-based measures among United States informal caregivers. Value Health. 2024;S1098301524001177. https://doi.org/10.1016/j.jval.2024.03.003.
7. Long C, Mao Z, Yang Z. A Head-to-Head Comparison of EQ-HWB and EQ-5D-5L in Patients, Carers, and General Public in China. Value Health. 2024;27(7):848–56. https://doi.org/10.1016/j.jval.2024.02.012.
8. 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. Qual Life Res. 2025. https://doi.org/10.1007/s11136-025-04061-3.
9. 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 Health. 2024;S1098301524000329. https://doi.org/10.1016/j.jval.2024.01.004.
10. 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 Health. 2025;S1098301525025124. https://doi.org/10.1016/j.jval.2025.07.028.
11. 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 Health. 2025;28(3):449–59. https://doi.org/10.1016/j.jval.2024.12.003.
12. Mukuria C, Peasgood T, McDool E, Norman R, Rowen D, Brazier J. Valuing the EQ Health and Wellbeing Short Using Time Trade-Off and a Discrete Choice Experiment: A Feasibility Study. Value Health. 2023;26(7):1073–84. https://doi.org/10.1016/j.jval.2023.02.008.
13. Peasgood T, Mukuria C, Brazier J, Kreimeier MO, Luo SN, et al. Developing a New Generic Health and Wellbeing Measure: Psychometric Survey Results for the EQ-HWB. Value Health. 2022;25(4):525–33. https://doi.org/10.1016/j.jval.2021.11.1361.
14. Vos T, Allen C, Arora M, Barber RM, Bhutta ZA, Brown A, et al. Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. Lancet. 2016;388(10053):1545–602. https://doi.org/10.1016/S0140-6736(16)31678-6.
15. Firth J, Torous J, Nicholas J, Carney R, Pratap A, Rosenbaum S, et al. The efficacy of smartphone-based mental health interventions for depressive symptoms: a meta-analysis of randomized controlled trials. World Psychiatry. 2017;16(3):287–98. https://doi.org/10.1002/wps.20472.
16. Linardon J, Messer M, Goldberg SB, Fuller-Tyszkiewicz M. The efficacy of mindfulness apps on symptoms of depression and anxiety: An updated meta-analysis of randomized controlled trials. Clin Psychol Rev. 2024;107:102370. https://doi.org/10.1016/j.cpr.2023.102370.
17. Van Dam NT, Targett J, Burger A, Davies JN, Galante J. Development and validation of the Inventory of Meditation Experiences (IME) [Internet]. 2024 Jan 30. https://doi.org/10.31234/osf.io/3ht95.
18. Peasgood T, Mukuria C, Carlton J, Connell J, Devlin N, Jones K, 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? Eur J Health Econ. 2021;22(7):1067–81. https://doi.org/10.1007/s10198-021-01306-z. PubMed PMID: 33909157.
19. Kessler RC, Barker PR, Colpe LJ, Epstein JF, Gfroerer JC, Hiripi E, et al. Screening for serious mental illness in the general population. Arch Gen Psychiatry. 2003;60(2):184–9. https://doi.org/10.1001/archpsyc.60.2.184 PubMed PMID: 12578436.
20. Cornelius BL, Groothoff JW, van der Klink JJ, Brouwer S. The performance of the K10, K6 and GHQ-12 to screen for present state DSM-IV disorders among disability claimants. BMC Public Health. 2013;13:128. https://doi.org/10.1186/1471-2458-13-128. PubMed PMID: 23402478; PubMed Central PMCID: PMC3575398.
21. Tennant R, Hiller L, Fishwick R, Platt S, Joseph S, Weich S, et al. The Warwick-Edinburgh Mental Well-being Scale (WEMWBS): development and UK validation. Health Qual Life Outcomes. 2007;5(1):63. https://doi.org/10.1186/1477-7525-5-63.
22. Stewart-Brown S, Tennant A, Tennant R, Platt S, Parkinson J, Weich S. Internal construct validity of the Warwick-Edinburgh Mental Well-being Scale (WEMWBS): a Rasch analysis using data from the Scottish Health Education Population Survey. Health Qual Life Outcomes. 2009;7(1):15. https://doi.org/10.1186/1477-7525-7-15.
23. Ng Fat L, Scholes S, Boniface S, Mindell J, Stewart-Brown S. Evaluating and establishing national norms for mental wellbeing using the short Warwick–Edinburgh Mental Well-being Scale (SWEMWBS): findings from the Health Survey for England. Qual Life Res. 2017;26(5):1129–44. https://doi.org/10.1007/s11136-016-1454-8. PubMed PMID: 27853963; PubMed Central PMCID: PMC5376387.
24. Cohen S, Kamarck T, Mermelstein R. A Global Measure of Perceived Stress. J Health Soc Behav. 1983;24(4):385–96. https://doi.org/10.2307/2136404.
25. Malik AO, Peri-Okonny P, Gosch K, Thomas M, Mena C, Hiatt WR, et al. Association of Perceived Stress Levels With Long-term Mortality in Patients With Peripheral Artery Disease. JAMA Netw Open. 2020;3(6):e208741. https://doi.org/10.1001/jamanetworkopen.2020.8741. PubMed PMID: 32573710; PubMed Central PMCID: PMC7312389.
26. Cheung F, Lucas RE. Assessing the validity of single-item life satisfaction measures: results from three large samples. Qual Life Res. 2014;23(10):2809–18. https://doi.org/10.1007/s11136-014-0726-4.
27. Caci H. CORTESTI: Stata module to test equality of two correlation coefficients [Internet]. Boston College Department of Economics; 2000 [cited 2026 Jun 24]. (Statistical Software Components No. S407302). Available from: https://ideas.repec.org/c/boc/bocode/s407302.html.
28. Cohen J. Quantitative methods in psychology: A power primer. Psychol Bull. 1992;112:1155–9.
29. 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 Health. 2025;28(8):1241–52. https://doi.org/10.1016/j.jval.2025.04.2151.
30. Australian Bureau of Statistics. Mental health- Healthy throughout life [Internet]. ABS. 2025 [cited 2025 Jan 9]. Available from: https://www.abs.gov.au/statistics/measuring-what-matters/measuring-what-matters-themes-and-indicators/healthy/mental-health?utm_source
31. Devlin N, Parkin D, Janssen B. Methods for Analysing and Reporting EQ-5D Data [Internet]. Cham: Springer International Publishing; 2020 [cited 2023 Aug 21]. Available from: https://link.springer.com/book/10.1007/978-3-030-47622-9.
32. Feng YS, Kohlmann T, Peasgood T, Engel L, Mulhern B, Pickard AS. Scoring the EQ-HWB-S: can we do it without value sets? A non-parametric item response theory analysis. Qual Life Res. 2024;33(5):1211–22. https://doi.org/10.1007/s11136-024-03601-7.
33. Adams J, Davies J, Wattanatakulchat P, Galante J, D'Alfonso S, Van Dam NT. Real-world meditation app engagement: longitudinal study of the medito meditation app. JMIR Mhealth Uhealth. 2026;14:e79366. https://doi.org/10.2196/79366.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.