Measuring mental well-being in Denmark: Validation of the original and short version of the Warwick-Edinburgh mental well-being scale (WEMWBS and SWEMWBS) and cross-cultural comparison across four European settings

Koushede, Vibeke; Lasgaard, Mathias; Hinrichsen, Carsten; Meilstrup, Charlotte Riebeling; Nielsen, Line; Rayce, Signe Lynne Boe; Torres-Sahli, Manuel; Gudmundsdottir, Dora Gudrun; Stewart-Brown, Sarah; Santini, Ziggi Ivan

Published in:
Psychiatry Research

DOI:
10.1016/j.psychres.2018.12.003

Publication date:
2019

Document version
Publisher's PDF, also known as Version of record

Document license:
CC BY-NC-ND

Citation for published version (APA):
Measuring mental well-being in Denmark: Validation of the original and short version of the Warwick-Edinburgh mental well-being scale (WEMWBS and SWEMWBS) and cross-cultural comparison across four European settings

Vibeke Koushede, Mathias Lasgaard, Carsten Hinrichsen, Charlotte Meilstrup, Line Nielsen, Signe Boe Rayce, Manuel Torres-Sahli, Dora Gudrun Gudmundsdottir, Sarah Stewart-Brown, Ziggi Ivan Santini

ARTICLE INFO

Keywords:
- Mental well-being
- Mental health
- Public health
- Measurement

ABSTRACT

This study examined the psychometric properties of the Danish WEMWBS and its short version (SWEMWBS) in a Danish population sample, and compared scores in Denmark with scores representative of three other European settings. A total of 3,508 Danish men and women aged 16–95 filled out an electronic survey. Face validity was examined by cognitive interviews. Content validity was assessed by examining response distributions and construct validity by confirmatory factor analysis, measurement invariance, and relations to other similar measures. Overall mental well-being scores were calculated, as well as stratified by sex and age. Support was found for the single-factor hypothesis, yielding good model fits for both versions of the scale. Both scales have high internal consistency. Correlations with mental health measures were largely in line with expectations. The highest mental well-being scores were reported for Catalonia, followed by Denmark, Iceland, and England. The (S)WEMWBS appear to be appropriate instruments to measure mental well-being in the Danish population. The present findings encourage the use of the scales, particularly SWEMWBS, in epidemiological, intervention and evaluation studies in research and practice. Cross-cultural comparisons like the one reported here may be essential to inform international mental health policy.

1. Introduction

Mental well-being is a concept regarded as encompassing dimensions of hedonic (positive feelings, affect, emotions) and eudemonic (positive functioning, mindset and relationships) well-being (Stewart-Brown, 2013). Over the past decade, interest in the concept of mental well-being has increased along with the recognition of its impact on public health (WHO, 2013). This has led to an increased interest among researchers, policymakers and service providers to improve knowledge about mental well-being measurement and application (Forsman et al., 2015; Huppert, 2014). In spite of mental well-being being closely connected to healthy life years and productivity, how to promote mental well-being has generally been under-researched, partly due to a lack of appropriate population-based measures (Huppert, 2014; Huppert and Whittington, 2003).

The Warwick-Edinburgh Mental Well-being Scale (WEMWBS) was developed in the United Kingdom (UK) to meet the need for a psychometrically robust measure that would enable monitoring of mental well-being in the general population and the evaluation of projects, programs and policies which aim to improve mental well-being (Stewart-Brown, 2015a,b; Tennant et al., 2007). Mental well-being is desirable in its own right and is also a modifiable determinant of...
longevity and a protective factor against future disease (Chida and Steptoe, 2008; Siahpush et al., 2008). The scale covers both feeling and functioning aspects of mental well-being which it represents as the positive end of a continuum of mental health. Licenses to use WEMWBS are provided free on application.

The scale is now widely used in the UK and other parts of Europe, where it is considered an appropriate tool to measure mental well-being (Stewart-Brown, 2015a,b). The WEMWBS has been validated in various populations and among different subgroups including adolescents, clinical samples and ethnic minority samples (McKay and Andretta, 2017; Smith et al., 2017; Stewart-Brown, 2015a,b; Stewart-Brown et al., 2011; Trousselard et al., 2016). The scale has been translated into more than 25 languages including Arabic, Urdu, Japanese and Chinese (Taggart, 2015), and validated in e.g. Norwegian, Swedish, Italian, Dutch, German, French and Spanish (Castelví et al., 2014; Forero et al., 2014; Haver et al., 2015; Smith et al., 2017; Taggart, 2015). Previous validation studies have found that the WEMWBS is considered easy to complete and that it provides a credible measure of mental well-being (Stewart-Brown, 2015a,b). The scale has also been found to be sensitive to the changes which occur in the context of a broad variety of well-being promotion initiatives (Stewart-Brown, 2015a,b).

Some validation studies suggest that item redundancy might be an issue for the WEMWBS leading to the development of the 7-item short WEMWBS (SWEMWBS) (Stewart-Brown et al., 2009). The SWEMWBS is preferred by some researchers in terms of its psychometric properties and convenience for monitoring mental well-being. However, the seven items in the SWEMWBS relate more to functioning than to feeling and therefore offer a slightly different perspective on mental well-being (Stewart-Brown, 2015a,b). When content coverage is an issue, such as in research projects where for example the project or intervention aims to help participants develop their understanding of the concept of mental well-being, gathering data on the full 14-item scale may therefore be preferred (Stewart-Brown et al., 2009).

The (S)WEMWBS is a potentially promising scale for measuring mental well-being in Danish population studies. However, the (S)WEMWBS has not been validated in a Danish context. Validation of a scale in the national context in which it is meant to be used is important in order to ensure its appropriateness in the specific setting. If validity of the (S)WEMWBS can be ensured in a Danish setting, it would be of immense value as it would: a) provide a tool suitable for measuring mental well-being in population-based samples - in which there is currently a high demand; b) enable mental health promotion practitioners to evaluate their programmes using a practical mental well-being measure; and c) to conduct research on the distributions and predictors of mental well-being in order to inform experts and stakeholders relevant to national and international public mental health policy.

1.1. Aim

The aim of this study is to examine the psychometric properties and validate the original and the short Warwick-Edinburgh Mental Well-Being Scale among a Danish population sample, as well as comparing mental well-being scores in Denmark with scores representative of three other European settings.

2. Methods

2.1. Study design

We used data from a national cross-sectional survey The Danish Mental Health and Well-being Survey 2016 (Nielsen et al., 2017). The survey was carried out by Statistics Denmark. A random representative sample of Danish men and women aged 16 years and above was drawn from the Danish Civil Registration System. Statistics Denmark sent an electronic letter to the sampled individuals in October 2016 with information about the study and an invitation to participate. After a week a reminder letter was sent, and after yet another week a final reminder was sent.

2.2. Sampling

In total 5,050 men and 5,200 women were contacted. Out of them 1,656 men and 1,852 women responded to the web-based survey resulting in a response rate of 33% for men and 36% for women. In terms of non-response, 5,854 did not respond to the invitation to participate, 463 only partially completed the survey, 183 refused to participate, 3 could not participate due to language barriers, 213 could not participate due to privacy protection, and 26 could not participate due to medical conditions or disability.

2.3. Ethics

There is no formal agency for ethical approval of questionnaire-based survey studies in Denmark. The study complies with the Helsinki 2 declaration on ethics and is registered with the Danish Data Protection Authority, and confidentiality and privacy requirements were met. The participants’ voluntary completion and return of the survey questionnaires constituted implied consent.

2.4. Measures

All measures included in this study were self-administered.

2.4.1 (S)WEMWBS: The WEMWBS is a 14-item scale with 5 response categories, with a total score ranging from 14–70. The SWEMWBS consists of 7 items with a total score ranging from 7–35 (see table 2). The items are all positively worded. Respondents are required to describe their experience of each statement over the past two weeks using a 5-point Likert scale (‘none of the time’, ‘rarely’, ‘some of the time’, ‘often’, ‘all of the time’). The overall WEMWBS score is calculated by summing the scores for each item, with equal weights. A higher WEMWBS score therefore indicates a higher level of mental well-being (Tennant et al., 2007). The SWEMWBS is scored similarly by creating a sum score for all the items with final scores transformed to enhance scaling properties (available online) (Stewart-Brown, 2015a,b). Participants were asked to respond to the 14-item scale, of which the same responses were used for the 7-item scale, i.e. participants were not asked to respond to the seven same items twice.

The scale was translated into Danish in line with recommendations set forth by the World Health Organization for the translation and adaptation of scales, which includes 1) forward-translation, 2) expert panel back-translation, 3) pre-testing and cognitive interviewing, and 4) final version. The details of the translation methodology can be found elsewhere (WHO, 2018).

2.4.2. Other measures

Five additional measures were included in this study to assess relations to other variables. The measures were chosen to include similar concepts to the (S)WEMWBS and concepts expected to be associated with mental well-being. The measures included for analysis were:

2.4.2.1. WHO-5 (Topp et al., 2015) covering overall well-being: 5 items given a score from 0–5, then added up and multiplied by 4, and scored into a continuous scale 0–100, with high scores indicating high levels of mental well-being.

2.4.2.2. Self-rated health: Single item for self-rated health which asks respondents to rate their overall health (physical as well as mental), five response categories from poor to excellent ranging from 1–5, higher scores indicate better self-rated health. The question is part of the
Danish national health and morbidity surveys (Christensen et al., 2014).

2.4.2.3. Discomfort and pain: Six items for symptoms of discomfort and pain within the past two weeks in: Shoulder or neck; Back or lower back; Arms, hands, legs, knees, hips or joints; Headache; Stomach-ache; Difficulties sleeping. Each item coded 0 = symptom not present, 1 = symptom present. Items added up to a scale ranging from 0–6, with higher scores indicating a higher number of symptoms. The questions are part of the Danish National Health and Morbidity surveys (Christensen et al., 2014).

2.4.2.4. The Perceived Stress Scale (PSS) (Cohen et al., 1983) covering perceived stress and coping: 10 items each given a score from 0–4. Positive items reversed, and added up into a scale ranging from 0–40, with higher scores indicating higher levels of perceived stress.

2.4.2.5. The Patient Health Questionnaire for Depression and Anxiety (PHQ-4): Data on poor mental health was collected using the PHQ-4 which asks participants about their experience of core depressive and anxiety symptoms over the past two weeks as specified by DSM-IV (Kroenke et al., 2009). There are 4 items for depression/anxiety, each item is given a score from 0–3, then scored into continuous scale 0–12, with higher scores indicating high level of depression/anxiety.

Other variables included in the present study were: sex, age, education, and employment.

2.4.3. Steps of validation and statistical procedures

Validation of the scales examined: 1) the face validity, 2) construct validity assessing goodness of fit and measurement invariance through confirmatory factor analysis, as well as relations to other or similar measures, 3) content validity including distribution of responses and potential floor and ceiling effects, 4) assessment of internal consistency. Quantitative analyses were completed using the statistical package R (R. 2018), and the lavaan package for structural equation modelling (Rosseel, 2012).

2.4.3.1. Face validity. Cognitive interviewing techniques were used to examine the face validity of the scale (i.e. do people understand the questions in the way they were intended) and how participants process and respond to the scale. Eleven face-to-face interviews were conducted with six men and five women aged 20–77 years. Participants were selected striving for a variation on age, sex and education, attributes known to be associated with mental health and health literacy (Bo et al., 2014; Christensen et al., 2017). Interview data was analysed using the Framework Approach based on the four stages (comprehension, retrieval, judgement, and response) of Tourangeau et al.’s model for survey response (Tourangeau et al., 2000).

2.4.3.2. Construct validity. Confirmatory factor analysis (CFA) based unweighted least squares means and variance adjusted (ULSMV) was performed to assess goodness of fit and measurement invariance for a single-factor model of both the Danish WEMWBS and SWEMWBS. To enhance cross-validity of the analyses, two independent subsets were randomly created for assessing global fit (n = 1150) and measurement invariance (n = 2358). As recommended by Hoyle and Panter (1995), several fit indices were used including the Root Mean Square Error of Approximation (RMSEA), the Comparative Fit Index (CFI), and the Tucker-Lewis Index (TLI). Values greater than 0.95 for the CFI and TLI were considered to reflect good model fit. A RMSEA value of 0.06 or less is considered to indicate a good fit, although values up to 0.08 were considered acceptable (Hu and Bentler, 1999; Jöreskog and Sörbom, 1993). Measurement invariance was evaluated across sex (women vs. men) and age groups (16–54 years of age vs. 55+) performing a Likelihood-Ratio Test (LRT).

Criterion validity was assessed by calculating correlations between the WEMWBS scores and the SWEMWBS scores. Convergent validity was assessed by calculating correlations between the (S)WEMWBS and WHO-5. Discriminant validity was assessed by calculating correlations between the (S)WEMWBS and Self-rated health (SRH), education, symptoms of discomfort and pain, PSS, and PHQ-4. We hypothesized that the (S)WEMWBS scores would show a strong positive correlation with the positively phrased measure of general well-being (WHO-5) (Tennant et al., 2007), moderate positive association with SRH (McKay and Andretta, 2017), and moderate negative associations with scales measuring the negative aspects of physical or mental health status (symptoms of discomfort and pain, PSS and PHQ-4) (McKay and Andretta, 2017) (based on Cohen’s rule of thumb, i.e. small: r = 0.1; moderate = 0.3; large = 0.5, Cohen et al., 2003).

Based on the findings of recent Danish health and morbidity studies, we hypothesized that men would score higher than women on the (S) WEMWBS and that the scales would show a positive association with education (Christensen et al., 2014, 2017). The latter association was hypothesized to be mild to moderate based on recent studies suggesting that mental well-being is less sensitive to socioeconomic patterning compared to poor mental health (Nielsen et al., 2016). Differences in scores across sex and education were assessed using linear regression analysis.

2.4.3.3. Content validity. The distribution of responses highlighted the frequency of popular responses and any floor and ceiling effects. Total scores were examined for floor and ceiling effects. Instruments exhibit floor or ceiling effects if more than 15% of participants record the lowest or highest score (McHorney and Tarlov, 1995).

2.4.3.4. Internal consistency. Cronbach’s α and McDonald’s ω were calculated as reliability indices of the total scores. Internal consistency estimates of > 0.70 were sought (Baggozi and Yi, 2012).

2.4.3.5. Cross-cultural comparison of mental well-being scores. Total (S) WEMWBS scores were computed with weights applied to generate nationally representative estimates. The Danish mental well-being scores were reported along with scores based on data representative of three other European settings, specifically Iceland, England, UK, and Catalonia, Spain. Overall scores were reported as well as scores stratified by age and sex. Information regarding survey and sampling in Iceland, England, and Catalonia is provided in Appendix 1.

3. Results

3.1. Participant characteristics

319 (9.1%) were 16–25 years old, 735 (21.0%) were 26–44 years old, 1,437 (41.0%) were 45–64 years old, and 1,017 (28.9%) were 65–95. The mean age was 52.1 years. Among participants 2,528 (72.0%) were either married or living with a partner. 1,919 (54.7%) were employed.

3.2. Face validity

Participants found the scales easy to complete with no major problems in terms of understanding. All items are shown in Table 2. Item b) I have been feeling useful and item l) I have been feeling loved evoked a few comments on the ambiguity caused by the context sensitiveness of the questions (i.e. have I been feeling useful or loved at work, in school, in my family or as a citizen?). The concepts of feeling useful and loved were understood consistently among participants. The wording used for item n) (cheerful in Danish translated into fornøjet) was considered outdated or quirky. Nevertheless, participants felt able to provide an answer.
The current study tested the construct validity and cross-cultural comparison of mental well-being scores of the (S)WEMWBS in Denmark and other countries. The study included a large sample size with representatives from all social groups above the age of 16. Some limitations should be kept in mind when interpreting the results. Consistently, men reported slightly higher mental well-being scores than women in all four countries/regions. In terms of age differences, scores rose with age in northern settings, but fell with age in southern settings. Differences, scores rose with age in northern settings, but fell with age in southern settings. Differences, scores rose with age in northern settings, but fell with age in southern settings. Differences, scores rose with age in northern settings, but fell with age in southern settings. Differences, scores rose with age in northern settings, but fell with age in southern settings.
Convergent and discriminant validity tests were conducted to assess the relationship between the scales and other measures. The results indicated that the SWEMWBS displayed high internal consistency for both sex and age. The scales were shown to be useful for detecting overall improvement in mental well-being in the population. The SWEMWBS explained a substantial amount of the variance in WEMWBS scores, confirming item relationships with good model fit for both scales. The results from the convergent and discriminant validity tests suggest that the (S)WEMWBS share common features with the WHO-5, and are inversely related to the PSS and the PHQ-4, in line with the original validation study (Tennant et al., 2007). The (S)WEMWBS have proved very popular both in their country of origin and internationally (Shah, Steiner, Petrou, Johnson, and Stewart-Brown, in press) because they have enabled studies of risk and protective factors for mental well-being as well as intervention evaluation. Cognitive interviews with the general public confirm that the range of items in these measures offers a recognisable picture of a previously unfamiliar concept (Stewart-Brown, 2013). This is likely an important component of the scale's popularity in supporting research and evaluation in subtle ways that are hard to measure. The scales have proved responsive to change in evaluation studies (Maheswaran et al., 2012; Shah et al., Under review) and also in RCTs (Shah et al., in press). Studies of risk and protective factors suggest, surprisingly, that mental well-being is much less strongly associated with socio-demographic factors and health related lifestyles than mental illness (Ng Fat et al., 2016).
mentally well-being scores found for Catalonia compared to those reported in the aforementioned studies may be related to within-country variation, i.e. other studies reported on Spain as a whole rather than Catalonia in isolation. Catalonia is a region quite distinct from the rest of Spain, with considerable differences in ethnicity, cultural values and economic conditions, and it is possible that these differences may explain our findings rather than the properties of the different measures. That said, we were not able to identify any other studies indicating that mental wellbeing is considerably higher in Catalonia than in the rest of Spain. Finally, due to the different sources of survey data, we could not test for measurement invariance across the four settings, and we cannot rule out that this could also have had implications for the findings. Thus, explanations for the reasons behind the differences in ratings between our study and the aforementioned reports remain speculative.

Different measures capture different constructs. It is important to understand the differences between them and their underlying mechanisms in relation to these different measures in order to guide policy and practice most appropriately. Comparison of mental well-being scores across different European settings offer valuable insights that might otherwise have remained unknown, as well as allowing research into risk and protective factors that vary between countries. Such research is strongly needed to inform international mental health policy, and scales that are valid across cultures and languages are essential to support this research. Finally, multicultural validity testing is warranted to examine how (S)WEMWBS functions across different cultural settings.

4.1. Conclusion

The (S)WEMWBS are shown to be valid and appropriate instruments to measure mental well-being in the Danish population. The present findings encourage the use of these measures in epidemiological, intervention and evaluation studies, where they are likely to be valuable for research as well as practice. While the WEMWBS scale may be used, the short version explains a substantial amount of the original scale, and may be preferred for the sake of practicality. The highest mental well-being scores were reported for Catalonia, followed by Denmark, Iceland, and England. Our findings offer insights into distributions of mental well-being on a global level that would otherwise likely remain unknown. Future cross-national investigations are therefore strongly warranted, as they may be essential to inform international mental health policy.

Conflict of interest

The Authors declare that there is no conflict of interest.

Acknowledgements

We are grateful to Jordi Alonso and Gemma VIlagut from IMIM-Institut Hospital del Mar d'Investigacions, Me¨diques in Barcelona Spain for their comparative analysis of (S)WEMWBS scores in Catalonia; and Chinthana Perera from Division of Health Sciences, Warwick Medical School for analysis of the Health Survey for England.

Supplementary materials


Appendix

Data sources, designs, and samples.
Iceland

Data on wellbeing using the Short Warwick-Edinburgh Mental Well-Being Scale (SWEMWBS) were extracted from the survey Health and Wellbeing of Icelanders 2017, which was conducted in October 2017. Health and Wellbeing of Icelanders 2017 is a national questionnaire survey among 9887 randomly selected Icelanders, 18 years and older of which 6.776 (68.5%) participated and 3.644 (64.2%) completed SWEMWBS.

England, UK

Data on wellbeing using the Warwick-Edinburgh Mental Well-Being Scale (WEMWBS) were extracted from the Health Survey England 2016 (HSE). HSE 2016 survey sampled one dwelling unit from 9558 addresses randomly selected from 531 postcode sectors to provide a representative sample of adults (> 16 years) living in private households. All consenting adults living in the sample dwelling units were interviewed. 8011 adults were interviewed, 55% of those approached. Of these 7071 (88.3%) completed WEMWBS.

Catalonia, Spain

The Spanish version of the WEMWBS was administered to a representative sample of non-institutionalized general population 15 years or older in Catalonia, a northeastern region of Spain with about 7.5 million inhabitants, as part of the 2016 Catalan Health Interview Survey (CHIS) Waves 12 and 13. The sample was selected using a stratified multistage random sampling design. Computer-assisted personal interviews (CAPI) were carried out by trained interviewers at the respondents’ households from January 2016 to February 2017.

References


