

Validation of the German Warwick-Edinburgh Mental Well-Being Scale (WEMWBS) in a community-based sample of adults in Austria: a bi-factor modelling approach

Gert Lang^{1,2}  · Almut Bachinger^{1,3}

Received: 26 April 2016 / Accepted: 17 November 2016 / Published online: 3 December 2016
© Springer-Verlag Berlin Heidelberg 2016

Abstract

Aim This article reports on the validation of the German version of the Warwick-Edinburgh Mental Well-Being Scale (WEMWBS) in a community-based sample of adults living in Vienna, Austria ($n = 625$).

Subject and methods For psychometric property testing, survey data were rigorously analysed with advanced multivariate methods under aspects of validity and reliability. The properties of the total scale scores are examined by descriptive statistics.

Results There was a high response rate to the scale (94.2%) with uni-modal distributed and slightly skewed items. The one-factor Confirmatory Factor Analysis (CFA) model was not supported by the data, but the structure of the WEMWBS was excellently described by a bi-factor model [$\chi^2 = 160.63$ (60), RMSEA = 0.05, CFI = 0.97, SRMR = 0.03] with one strong general MWB factor (ECV = 0.70, $\omega_H = 0.81$) and three relatively weak residualised (group) factors (ECV < 0.15, $\omega_S < 0.29$), which were associated with the domains of MWB (positive affect, functioning and personal relationships). The mean total score of the 14-item scale was 54.5 (53.8–55.2) and slightly skewed, but with no floor or ceiling effects. There was a moderate correlation with the EUROHIS-

QOL scale ($r = 0.64$, $p < 0.01$) and score differences in major socio-demographic groups (age, education level, marital status).

Conclusions The German version of the WEMWBS achieved good validity and reliability in our test sample and is recommended for future applications. Advanced qualitative and quantitative research methods are suggested to further test the scale.

Keywords Mental health promotion · Mental well-being · Scale validation · Confirmatory factor analysis · Bi-factor analysis · Structural equation modelling

Abbreviations

$(\sum \lambda)^2$	Sum of loadings squared
$(\sum \lambda^2)$	Sum of squared loadings
CBA	Confirmatory Bi-factor Analysis
CFA	Confirmatory Factor Analysis
CFI	Comparative Fit Index
CI	Confidence interval
d.f.	Degrees of freedom
EBA	Exploratory Bi-factor Analysis
ECV	Explained Common Variance
GF	Group factor
m	Arithmetic mean
md	Median
mis	Missing answer
MWB	Mental Well-Being
N	Valid cases
ns	Statistically not significant
p	p-value
PMH	Positive Mental Health
QOL	Quality of life
RMSEA	Root Mean Standard Error of Approximation
s^2	Standard deviation

✉ Gert Lang
gert.lang@goeg.at; <http://www.fgoe.org>

¹ Research Institute of the Red Cross, Nottendorfer Gasse 21, A-1030 Vienna, Austria

² Austrian Health Promotion Foundation, a business unit of Gesundheit Österreich GmbH, Aspernbrückengasse 2, A-1020 Vienna, Austria

³ International Centre for Migration Policy Development (ICMPD), Gonzagagasse 1, A-1010 Vienna, Austria

s^3	Skewness
s^4	Kurtosis
SRMR	Standardised Root Mean Square Residual
WEMWBS	Warwick-Edinburgh Mental Well-Being Scale
χ^2	Chi-square test statistic
ω_H	Coefficient omega hierarchical
ω_S	Coefficient omega specific

Introduction

In recent years, mental well-being (MWB) and positive mental health (PMH)¹ have been increasingly discussed and assigned more importance especially within the framework of health promotion. In the public health practice and policy arena, validated instruments measuring MWB outcomes are therefore in high demand.

From a theoretical perspective there is still a lack of consensus about the fundamental underlying concepts of MWB but at least there is agreement that it is a complex theoretical construct, encompassing more than just the absence of mental illness and one that has to integrate two meta-theoretical perspectives: namely hedonic and eudaimonic. As they imply rather distinct views of human nature and accordingly approaches to life and well-being, in theory and practice, they have been the subject of intense debate since the times of ancient Greece (Ryan and Deci 2001: 143). Nevertheless, they have been given new impetus in the course of the discussion about PMH and MWB.

The concept of hedonism refers to pleasure or happiness, and good life or well-being, which is interpreted as a maximum accumulation of such moments. Therefore, different approaches in hedonic psychology assess the experience of well-being within a pleasure-pain continuum. For instance, subjective well-being uses three components, namely: life satisfaction, presence of positive and absence of negative mood (together summarised as happiness) (Ibid.: 144). Often, the meta-theoretical starting point fits in with the expectancy value approach, which suggests that “*well-being is a function of expecting to attain (and ultimately attaining) the outcomes one values, whatever those might be*” (Ibid.: 145). Diener et al. (2009: 63) define subjective well-being as “*a person's cognitive and affective evaluations of his or her life*”.

In contrast, the eudaimonic concept derives from Aristotle, who considered hedonism a vulgar ideal and that the pursuit of happiness makes humans slavish followers of their desires. While the hedonic paradigm emphasises the internal affective and cognitive aspects of well-being, the eudemonic concept adds an external dimension (Vanhoutte and Nazroo 2013). Good life goes beyond subjectively felt needs and also takes

into account objectively valid needs, as Fromm (1978–79) put it. According to Waterman (1993), the eudaimonic conception of well-being calls upon people to live in accordance with their ‘daimon’, or true self. Under such circumstances one would feel intensely alive and authentic, in a state of “*personal expressiveness*”. This concept involves more activity, development and personal growth (also as a result of being challenged and exerting effort) than being relaxed and happy. Ryff and Keyes (1995) embraced eudaimonic aspects as psychological well-being, which comprises six distinct dimensions of human actualisation: autonomy, environmental mastery, personal growth, positive relations with others, purpose in life and self-acceptance.

To date there is a consensus that both concepts need to be taken into account to understand MWB. Thus, MWB is conceived as a multi-faceted construct that comprises both hedonic and eudaimonic elements. The issue that is still at stake regarding a concept of MWB is that these elements are both distinct and overlapping and at the same time converging and diverging (Ryan and Deci 2001: 148).

Consequently, this constitutes a challenge for designing a measuring instrument for well-being, several of which are currently available or under development (Vásquez and Hervás 2012). One is the Warwick-Edinburgh Mental Well-Being Scale (WEMWBS), developed by Stewart-Brown and colleagues at the University of Warwick Medical School. The WEMWBS² aims to provide a short instrument to measure PMH and MWB, which is meaningful and easily understandable for people from various social backgrounds (Stewart-Brown 2013: 134).

The WEMWBS

The Warwick-Edinburgh Mental Well-Being Scale (WEMWBS) is a 14-item Likert scale that covers “*both hedonic and eudaimonic aspects of mental health including positive affect (feelings of optimism, cheerfulness, relaxation), satisfying interpersonal relationships and positive functioning (energy, clear thinking, self-acceptance, personal development, competence and autonomy)*” (Tennant et al. 2007: 3).

Respondents to the scale items are asked to describe their experience over a two-week reference period for which an ordinal five-point frequency answer format, ranging from 1 = “*none of the time*” to 5 = “*all of the time*”, is employed. Based on the assumption of one single general MWB factor, the total score is computed as the sum of all items. If respondents answer all items, the theoretical range of the WEMWBS is from 14 to 70 scale points with a higher score indicating a higher level of MWB (Stewart-Brown and Janmohamed 2008).

¹ The concepts MWB and PMH are often used interchangeably. In the following report, we will refer to MWB.

² By adaption and shortening of the Affectometer 2 scale (Tennant et al. 2006).

The WEMWBS was originally developed in English. So far, the scale has been translated into a number of languages (cf. Warwick Medical School 2016). To date, several validation studies have been published, based on both smaller, convenient and community-based samples (e.g. community-dwelling individuals) or cross-sectional samples (e.g. representative of a national population) from the UK (Clarke et al. 2011; Lloyd and Devine 2012; Taggart et al. 2013; Tennant et al. 2007), Spain (Castellví et al. 2014; López et al. 2013), Italy (Gremigni and Stewart-Brown 2011), Turkey (Keldal 2015), Pakistan (Waqas et al. 2015), Australia (Hunter et al. 2015), South Korea (Kim et al. 2014), Brazil (Santos et al. 2015) and Argentina (Azcurra 2015). The English language version was used most frequently, followed by translated versions. The scale was most often tested with adult populations (i.e. 18 to 75 years of age), seldom with adolescents (or pupils, students, e.g. for pre-validation purposes) and only once specifically with older people (i.e. 71 to 96 years of age) (Azcurra 2015).

To date, psychometric studies examining the factor structure (construct validity) of the WEMWBS have substantial shortcomings. Although the hypothesised one-factor scale structure has been supported (tested primarily with exploratory and/or confirmatory factor analytic techniques), many of these studies yielded fit indices that were far from ideal (see also Hunter et al. 2015) when considering conventional criteria (Brown 2006). Some reported a quite good fit (Castellví et al. 2014; Clarke et al. 2011; Keldal 2015; Lloyd and Devine 2012; Tennant et al. 2007) but others reported a less than optimal fit of the empirical data to the hypothesised one-factor model (Gremigni and Stewart-Brown 2011; Hunter et al. 2015; Kim et al. 2014; López et al. 2013). In some applications a two or three factor model was considered, or even implemented (Azcurra 2015; Gremigni and Stewart-Brown 2011; López et al. 2013; Santos et al. 2015; Taggart et al. 2013; Waqas et al. 2015). In addition, adaptations to the scale were also implemented to improve the model fit, e.g. by the specification of residual error covariances (Clarke et al. 2011; Hunter et al. 2015) or even by a reduction of scale items (Gremigni and Stewart-Brown 2011; Santos et al. 2015).

Interestingly, studies have not sought to identify the sources of misfit in greater detail and have also failed to examine whether the scale measurement covers unique variance over and above a general factor of MWB and to test whether it allows for a precise scaling of individuals.

Reise et al. (2010) stated that the problem with traditional psychometric scale testing is that the wrong default model is often used. In particular for multi-faceted constructs and complex measures that include heterogeneous item content, research literature has pointed out that the item responses are possibly not only a reflection of the general factor (Chen et al. 2012; Reise 2012). Consequently, the assumption of uni-dimensionality is often too restrictive and can also result in model misfits. In contrast, a bi-factor modelling approach

provides a viable solution and alternative modelling strategy. A bi-factor model is more suitable for a factor structure of multi-faceted items that cover broad (i.e. hedonic and eudaimonic) concepts and highly related (i.e. relationships, affect, functioning) domains that are hypothesised to comprise a general (i.e. MWB) construct (cf. Chen et al. 2006: 190). Recently, bi-factor modelling has been widely implemented and has helped researchers clarify the measurement structure of scales in mental health-related contexts, too (Chen et al. 2013; de Bruin and du Plessis 2015; Deng et al. 2015; Jovanović 2015a, b; Longo et al. 2016; Mészáros et al. 2014). Specifically, Reise concluded that “(...) *bi-factor modeling potentially provides a solid foundation for conceptualizing psychological constructs, constructing measures, and evaluating a measure’s psychometric properties*” (Reise 2012: 667).

To the best of our knowledge, bi-factor modelling has not yet been applied in studies investigating the psychometric properties of the WEMWBS. The benefit of a bi-factor analysis is that it reveals information about residualised (group) factors above and beyond the hypothesised general MWB factor. The analysis allows a comparison of the strength of a general MWB factor relative to the group factors and estimates the size of the factor loadings of each item. In addition, it delivers additional insights for the appropriateness of the assumption of uni-dimensionality of the WEMWBS. If detected, a strong general factor where each item adequately loads the general factor would provide a justification for a WEMWBS total score.

Aims and objectives

The aim of the present research is to verify the theoretical assumptions of the WEMWBS. Therefore, the research questions are whether the theoretical construct of MWB—which is multi-faceted and comprises distinct and overlapping, but at the same time converging and diverging hedonic and eudaimonic elements—can be empirically proved by the WEMWBS and whether the scale is a valid measurement tool for MWB.

In order to achieve this aim an evaluation of the psychometric properties of the measurement scale was conducted with the first translated German version of the WEMWBS. This represents its preliminary validation, which was done in several steps: first, the face validity of the scale was assessed by a descriptive item analysis. Second, construct validity was tested by confirmatory factor analysis techniques for which a series of one-factor and bi-factor models were specified and compared. Third, criterion (concurrent) validity of the scale was assessed with another well-established external quality of life measure (the EUROHIS-QOL scale). Fourth, properties of the WEMWBS total score were evaluated. The analysis was of data from a community-based sample randomly drawn from adults of advanced age in Vienna, Austria.

Materials and methods

Survey instrument

After obtaining permission from the University of Warwick Medical School, the English WEMWBS version was translated by the authors of this article, based on the guidelines drawn up by Beaton and colleagues (Beaton et al. 2000), and aimed at achieving a cross-culturally sensitive adaptation and conceptually equivalent German version.

The translated scale (Table 1) was included in a questionnaire for an omnibus survey by the Viennese Red Cross, which included questions about the organisation's member magazine, opinions about charitable giving and general socio-demographic information. Respondents were additionally asked to take part in a supplementary survey about well-being and quality of life, which included the WEMWBS and EUROHIS-QOL scales.

Sampling and data collection

The gross population consisted of all supporting members of the Viennese Red Cross ($N = 77,191$ based on organisational data from 2013). The sampling frame comprised 41,684 members (54%) who lived in the community of Vienna and who had been members for at least 1 year (excluding staff). Out of this, a simple random sample of 6000 was drawn representing a fraction of 14.4% to whom the postal survey was administered in May 2014. The survey aims and the selection process were explained in an accompanying letter. The cover letter included information on how to complete and return the questionnaire in the enclosed envelope, free of charge. In addition,

it was possible to respond electronically by completing an identical online questionnaire. For ethical reasons, respondents were informed that participation was voluntary and data handling was based on confidentiality and anonymity. It was possible to contact the telephone service line for further information. In total, 625 people completed the questionnaire (only 14 responded online). After taking neutral non-responses (e.g. deceased, sick persons, unknown addresses) into account, the response rate to the survey was 10.4%. The data of the completed questionnaires were stored using SPSS statistical analysis software without personal identifiers.

Participants

In the sample, more than half of the respondents were females (58.4%) and around two fifths were male (41.6%); 63.5% were married or cohabiting with a partner; 40.1% had attended school for 10–12 years. The majority fell into higher age categories; 82.1% were 60 years or older and therefore 76.6% were retired. The majority had no migration background, but for 14.7%, at least one of their parents had been born abroad. The respondents' living area reflected a good representation of the Viennese population with regard to inner, outer and peripheral districts (wien.at 2014) (Table 2).

Analysis methods

Face, construct and criterion validity were analysed in several steps:

- (1) High readability and face validity are shown if there is a low rate of non-complete responses (Moser and Kalton

Table 1 English original and German translation of the 14-item WEMWBS

Item	Wording in English	Wording in German
v1	I've been feeling optimistic about the future	Ich habe mich in Bezug auf die Zukunft optimistisch gefühlt
v2	I've been feeling useful	Ich habe mich nützlich gefühlt
v3	I've been feeling relaxed	Ich habe mich entspannt gefühlt
v4	I've been feeling interested in other people	Ich habe mich für andere Menschen interessiert
v5	I've had energy to spare	Ich hatte viel Energie
v6	I've been dealing with problems well	Ich bin mit Problemen gut umgegangen
v7	I've been thinking clearly	Ich konnte klar denken
v8	I've been feeling good about myself	Ich habe mich wohl gefühlt
v9	I've been feeling close to other people	Ich habe mich anderen Menschen nahe gefühlt
v10	I've been feeling confident	Ich habe mich zuversichtlich gefühlt
v11	I've been able to make up my own mind about things	Ich war in der Lage, Entscheidungen zu treffen
v12	I've been feeling loved	Ich habe mich geliebt gefühlt
v13	I've been interested in new things	Ich habe mich für Neues interessiert
v14	I've been feeling cheerful	Ich habe mich fröhlich gefühlt

Source: Warwick-Edinburgh Mental Well-Being Scale (WEMWBS) © NHS Health Scotland, University of Warwick and University of Edinburgh

Table 2 Socio-demographic characteristics of survey participants (n = 625)

Variables	Categories	n	%
Total		625	100.0
Gender (n = 591)	Male	246	41.6
	Female	345	58.4
Age (n = 608)	Up to 49 years	39	6.4
	50–59 years	70	11.5
	60–69 years	158	26.0
	70–79 years	205	33.7
	80–89 years	111	18.3
	90 years or above	25	4.1
Civil status (n = 603)	Single, never married	71	11.8
	Married, cohabiting	312	63.5
	Divorced, separated	79	13.1
	Widowed	141	23.4
School attendance (n = 606)	Up to 9 years	103	17.0
	10–12 years	243	40.1
	13–15 years	158	26.1
	16 years or more	102	16.8
Employment status (n = 593)	Employed (full or part-time)	100	15.8
	Self-employed	17	2.7
	Retired	485	76.6
	Other ^a	31	5.0
Migration history/background ^b (n = 605)	No	513	84.8
	Yes	92	15.2
Living district in Vienna (n = 605)	Inner districts (1–9, 20)	161	26.6
	Outer districts (10–19)	310	51.2
	Peripheral districts (21–23)	134	22.2

^a For example, unemployed, in training or homemaker

^b Whether the responding person or at least one of their parents was born abroad

1971). Additionally, there should be no truncated distributions indicating floor/ceiling effects (Bortz and Döring 2006). Item (non-) responses were assessed by frequency analysis in SPSS and described by item mean (m), standard deviation (s^2), skewness (s^3) and kurtosis (s^4). It was expected that items have a meaningful variation around their means, with approximately normal ($s^3 < |2.0|$) and uni-modal ($s^4 = |7.0|$) distribution (Byrne 2012).

- (2) A uni-dimensional latent construct with fixed relations between MWB and the 14 items (v1–v14) was hypothesised (Brown 2006). To test construct validity (Cronbach and Meehl 1955), several models were tested: first, a Confirmatory Factor Analysis (CFA) with one general MWB factor with robust maximum likelihood (MLR) estimation was performed in Mplus 7 (Muthén and Muthén 2012). The CFA results were assessed by standardised factor loadings (> 0.50) and significant ($p < 0.050$) residuals and by plausible modification indices (Byrne 2012), e.g. error covariances of lower magnitude (Urban and Mayerl 2014). The models were

rigorously evaluated by goodness-of-fit statistics: χ^2 -test ($\chi^2/\text{d.f.} < 3.0$)³ (Kline 2011), Root Mean Standard Error of Approximation (RMSEA < 0.08), Comparative Fit Index (CFI > 0.95) and the Standardised Root Mean Square Residual (SRMR < 0.05) (Browne and Cudeck 1993; Hu and Bentler 1999).

Second, for bi-factor modelling Reise (2012: 677) emphasised the necessity of an exploratory inspection of the data before carrying out a Confirmatory Bi-factor Analysis (CBA) in which additional group factors may be detected over and above the general MWB factor. In any set of items, the CBA enables comparison of the strength of a general factor relative to group factors. Besides the overall model fit criteria stated above, the CBA was evaluated by the proportion of total variance in items explained by the general factor and group factors. For reliability assessment, general MWB factor

³ Note: For a good fit, the χ^2 value should be non-significant, but is rarely obtained when the sample is large, as is the case in the present study.

saturation is expressed by McDonald's coefficient omega hierarchical ($\omega_H > 0.80$); the specific factor saturation of the group factors is evaluated by the coefficient omega specific (ω_S) (McDonald 1999; Zinbarg et al. 2005). Furthermore, if there is a relatively high proportion of common variance (ECV, e.g. > 0.70) explained by the general factor, this provides information about the validity of the uni-dimensionality of the WEMWBS (a high ECV for the group factors indicates multi-dimensionality). The higher the ECV, the stronger the general factor relative to the group factors.

- (3) Criterion validity (concurrent validity) was tested with a well-established measurement scale, the EUROHIS-QOL (Schmidt et al. 2006), derived from the WHO Quality of Life assessment tool (WHOQOL-100) (Power et al. 1998). It is a Likert scale self-assessment instrument of generic quality of life (QOL), with an ordinal five-point answer format covering a two-week reference period. The scale revealed good psychometric properties, especially for the German language version, e.g. high internal consistency ($\alpha = 0.80$) and factorial validity (CFI = 0.94, RMSEA = 0.08) (Brähler et al. 2007). Since previous studies revealed rather moderate correlations with various life satisfaction scales (Kim et al. 2014; López et al. 2013; Tennant et al. 2007), a correlation of $r > 0.50$ between the two scales was hypothesised.
- (4) Finally, the scale properties were analysed by the mean (m) and 95% confidence interval (CI) of the total scale score across major socio-demographic groups. Similar to previous studies, no significant gender WEMWBS differences were expected but higher total scale scores were revealed for younger age groups and better-educated employed people living with a partner (married, cohabiting) and with no migration background/history.

Results

Face validity

Out of 625 respondents, 24 did not complete the supplement that included the WEMWBS and were consequently excluded from the analysis; 556 out of 601 responded to all 14 items of the scale, indicating a high response rate of 94.2%. The highest item non-response was 2.2% (item 3). All categories were used and frequency distributions mostly showed "often" as the modal answer category. All items had a mean between $m = 3.0$ (items v3, v5) and 4.0 (items v11-13), mostly with a standard deviation between $s^2 = 0.7$ (items v6-7, v11) and 0.9 (items v1-v3, v5, v12). Two items had a higher negative skew of -1.5 (item v11) and -1.2 (item v7) with a higher kurtosis ($s^4 > 1.7$) (Table 3).

Construct validity

First, the CFA tested general MWB in a one-factor model that showed significant and high factor loadings between 0.47 and 0.78. However, this model (1a) was not able to reproduce all the item variability as was shown by a bad overall fit: $\chi^2 = 593.10$ (77), RMSEA = 0.11, CFI = 0.82 and SRMR = 0.07. Eleven error covariances between pairs of items needed to be specified in model 1b to achieve a reasonable fit: $\chi^2 = 226.98$ (66), RMSEA = 0.06, CFI = 0.94, SRMR = 0.04 (Table 4).

Second, the exploratory bi-factor analysis in model 2a with three group factors achieved an acceptable fit: $\chi^2 = 247.62$ (52), RMSEA = 0.08, CFI = 0.93, SRMR = 0.03. The loading pattern indicated that the items mostly assemble according to the affective (GF1), functioning (GF2) and relationship (GF3) domains of MWB. Based on these results, a CBA in model 2b with three group factors (GF1-3) was specified (see item groupings in Table 5). Additionally, three significant ($p < 0.010$) error covariances between items v1 and v2 ($r = 0.33$), v1 and v10 ($r = 0.29$) and v9 and v12 ($r = 0.21$) were defined, which resulted in an excellent overall fit of the model: $\chi^2 = 160.63$ (60), RMSEA = 0.05, CFI = 0.97, SRMR = 0.03.

In a model comparison, CBA model 2b was the best fitting model and demonstrated superiority. It showed a strong general MWB factor, with significant ($p < 0.001$) and high factor loadings between 0.51 (v7) and 0.70 (v10) on the general MWB factor. All items except v7 had higher loadings on the general factor than on the group factors, indicating that the variances were predominantly explained by the general MWB factor. Also noteworthy was the finding that no factor loading of group factor GF3 was significant ($p > 0.050$), indicating that this group factor may not exist above and beyond the general MWB factor and that the variance related to these items is already explained by the general MWB factor⁴ (Table 5).

The general MWB factor explained 70% of the common variance (ECV), while the group factors explained only a small proportion between 6 and 14%, which illustrates that the general factor accounted for a substantially greater amount of variance than the three group factors.

Superior to Cronbach's alpha⁵ (Zinbarg et al. 2005), McDonald's coefficient omega hierarchical (ω_H) is used in bi-factor models as a measure of reliability and indicates how precisely the WEMWBS items measure the general factor of MWB. In this study, ω_H was 0.81 and showed that the greatest

⁴ A more parsimonious but incomplete CBA model including only group factors GF1 and GF2 but without GF3 also resulted in an excellent model fit: $\chi^2 = 186.54$ (64), RMSEA = 0.06, CFI = 0.96, SRMR = 0.03.

⁵ Cronbach's alpha (α) was 0.92 for the total score and 0.74, 0.66 and 0.60 for the group factors.

Table 3 Frequency distribution and descriptive statistics of the scale items

Item	Short item wording	n	mis (%)	Relative frequencies (%) ^a					m	s ²	s ³	s ⁴
				1	2	3	4	5				
v1	Optimistic	595	1.0	2.2	13.6	28.9	39.3	16.0	3.53	0.99	-0.36	-0.42
v2	Useful	595	1.0	1.7	5.9	23.0	45.7	23.7	3.84	0.91	-0.68	0.38
v3	Relaxed	588	2.2	1.5	14.5	31.5	43.9	8.7	3.44	0.90	-0.37	-0.29
v4	Interested in people	597	0.6	0.3	4.5	20.3	43.6	31.3	4.01	0.85	-0.59	-0.10
v5	Energy to spare	594	1.2	1.3	17.0	29.8	40.2	11.6	3.44	0.95	-0.25	-0.59
v6	Dealing with problems	594	1.2	0.2	3.2	22.7	51.2	22.7	3.93	0.77	-0.39	-0.04
v7	Thinking clearly	598	0.5	0.3	1.2	7.9	38.5	52.2	4.41	0.71	-1.20	1.77
v8	Good about myself	594	1.2	0.7	6.6	21.5	52.5	18.7	3.82	0.83	-0.62	0.36
v9	Close to people	594	1.2	0.8	7.1	19.9	47.6	24.6	3.88	0.89	-0.66	0.18
v10	Confident	594	1.2	0.7	8.4	25.1	46.1	19.7	3.76	0.89	-0.46	-0.17
v11	Able to make up mind	595	1.0	0.5	1.3	6.6	34.1	57.5	4.47	0.72	-1.51	2.90
v12	Loved	590	1.8	1.7	6.4	19.8	33.2	38.8	4.01	1.00	-0.82	0.05
v13	Interested in things	598	0.5	0.2	3.7	13.5	38.3	44.3	4.23	0.83	-0.91	0.30
v14	Cheerful	596	0.8	0.3	9.9	26.8	45.3	17.6	3.70	0.88	-0.36	-0.41

n valid cases, mis missing answer, m arithmetic mean, s² standard deviation, s³ skewness, s⁴ kurtosis

^aThe response scale for all items was: 1 = none of the time, 2 = rarely, 3 = some of the time, 4 = often, 5 = all of the time

amount of variance of the composite score is attributable to the general factor. The omega specific (ω_S) coefficients were 0.26, 0.28 and 0.09, which showed poor subscale reliability.

Criterion validity

As hypothesised, the Spearman’s rank correlation between the EUROHIS-QOL and the WEMWBS was moderate with $r = 0.64$ ($p < 0.010$). As expected, the total score was not different between genders but discriminated between age groups. In addition, the mean score was higher for those living with a partner (married, cohabiting) compared to those with no partner (single, separated, divorced or widowed). People who attended school for less than 10 years scored significantly lower than those with more

years in school. There was no difference between occupational status and people with a migration background. Numerical details are shown in Table 6.

Scale properties

The total WEMWBS score ranged empirically between 25 and 70 with a mean value of $m = 54.5$ and a median of $md = 55.0$. There was a meaningful standard deviation $s^2 = 8.1$, and the 95% confidence interval was between 53.8 and 55.2. The shape of the distribution was slightly skewed ($s^3 = -0.40$, $s^4 = -0.21$). There were no floor/ceiling effects because only 0.2% had the lowest and 1.4% had the highest total score (Fig. 1).

Table 4 Model comparison

Model	χ^2	d.f.	RMSEA	90% CI		CFI	SRMR
				low	high		
One-factor							
1a. One-factor CFA	593.10	77	0.11	0.10	0.11	0.82	0.07
1b. One-factor CFA with 11 error covariances	226.98	66	0.06	0.06	0.07	0.94	0.04
Bi-factor							
2a. EBA with 1 general and 3 group factors	247.62	52	0.08	0.07	0.09	0.93	0.03
2b. CBA with 1 general, 3 group factors and 3 error covariances	160.63	60	0.05	0.04	0.06	0.97	0.03

CFA Confirmatory Factor Analysis, EBA Exploratory Bi-factor Analysis, CBA Confirmatory Bi-factor Analysis, χ^2 chi-square test statistic, d.f. degrees of freedom, RMSEA Root Mean Standard Error of Approximation, CI confidence interval, CFI Comparative Fit Index, SRMR Standardised Root Mean Square Residual

Table 5 Standardised factor loadings of one-factor and bi-factor models of the WEMWBS

Item	Short item wording	One factor	Bi-factor			
			General factor	GF1	GF2	GF3
v1	Optimistic	0.64	0.53	0.32		
v3	Relaxed	0.65	0.53	0.48		
v5	Energy to spare	0.71	0.67	0.27		
v8	Good about myself	0.76	0.62	0.58		
v10	Confident	0.78	0.70	0.30		
v14	Cheerful	0.77	0.67	0.47		
v6	Dealing with problems	0.62	0.60		0.39	
v7	Thinking clearly	0.52	0.51		0.55	
v11	Able to make up mind	0.57	0.59		0.49	
v13	Interested in things	0.54	0.55		0.22	
v2	Useful	0.61	0.66			0.06 ns
v4	Interested in other people	0.47	0.52			0.60 ns
v9	Close to other people	0.58	0.60			0.26 ns
v12	Loved	0.61	0.64			0.05 ns
	Sum of squared loadings ($\sum \lambda^2$)	5.68	5.06	1.05	0.74	0.43
	Sum of loadings squared ($\sum \lambda$) ²	77.95	70.16	5.84	2.70	0.76
	Omega hierarchical (ω_H)		0.81			
	Omega specific (ω_S)			0.26	0.28	0.09
	Explained common variance (ECV)		0.70	0.14	0.10	0.06

$n = 601$, GF group factor. All factor loadings were significant ($p < 0.001$) except those in italics (*ns*)

Discussion and conclusions

Discussion

The aim of the present research was to verify the theoretical assumptions of the German version of the WEMWBS by evaluating its validity and to assess whether it is a valid tool for the assessment of MWB. In contrast to the rather strong empirical assumption of the original authors of the scale (Tennant et al. 2007), the one-factor structure of the WEMWBS was not supported in the present study. In a non-parsimonious approach to model fitting, the one-factor model was only acceptable after the specification of 11 error covariances between items. Besides the varying cultural/country contexts, diverse study populations and language forms, this might also explain why the one-factor solution has not been able to achieve an excellent model fit in many previous studies.

By contrast, the CBA model indicated that the structure of the WEMWBS was best described by one general MWB factor and three group factors, which yielded the best fit for the data in the present sample. This demonstrated that the specific domains of MWB have unique explanatory power over and above general MWB. It is noteworthy that the bi-factor result primarily yielded strong support for the general MWB factor within the WEMWBS items, but gave only limited evidence for a viable multi-dimensional structure. All items had high

and significant loadings on the general factor and—*very importantly*—relatively low loadings on the group factors.

The bi-factor analysis revealed that the WEMWBS measures a strong general MWB factor and relatively weak residualised group factors that correspond to domains of MWB. It indicates that items measuring eudemonic (positive affect) and hedonic concepts (positive functioning, satisfying relationships) are indeed manifestations of the general MWB factor. It was empirically demonstrated that these domains do not reflect additional factors of MWB because of the small proportions of total and common variance explained by the residualised group factors as well as by their low reliabilities. A substantial amount of variance was accounted for by the general factor with high score reliability. As evidenced, the general factor accounted for 81% of the total test variance and 70% of the common variance. Although the discovery of the group factors suggests that they capture different domains of MWB over and above general MWB, the group factors accounted for only 9% of its observed variance and 6% of its common variance. Moreover, the low factor loadings indicated that group factors do not yield precise enough measures of unique aspects of MWB to be useful in practical applications. In other words, within the WEMWBS there is little common variance beyond general MWB, which suggests exclusively calculating a total WEMWBS score is both

Table 6 WEMWBS total scores across socio-demographic groups

Variable	Category	WEMWBS total score			
		n	m	95% CI	p
Total		566	54.5	53.8–55.2	
Gender	Male	226	54.7	53.7–55.7	<i>ns</i>
	Female	311	54.5	53.6–55.5	
Age	<59 years	100	54.6	53.1–56.1	**
	60–79 years	338	55.9	55.1–56.7	
	80 years or more	113	50.9	49.3–52.6	
Partner	Single, separated, divorced, widowed	254	53.8	52.7–54.9	*
	Married, cohabiting	295	55.4	54.6–56.3	
Education	<10 years	86	52.4	50.5–54.4	*
	10–12 years	224	55.2	54.2–56.2	
	13–15 years	143	54.6	53.3–55.9	
	16 years or more	96	55.4	54.0–55.3	
Occupation	Employed	103	55.2	53.7–56.6	<i>ns</i>
	Retired	432	54.5	53.8–55.3	
	Other	14	55.4	50.6–60.3	
Migration	Yes	83	54.4	52.7–56.1	<i>ns</i>
	No	465	54.7	54.0–55.4	

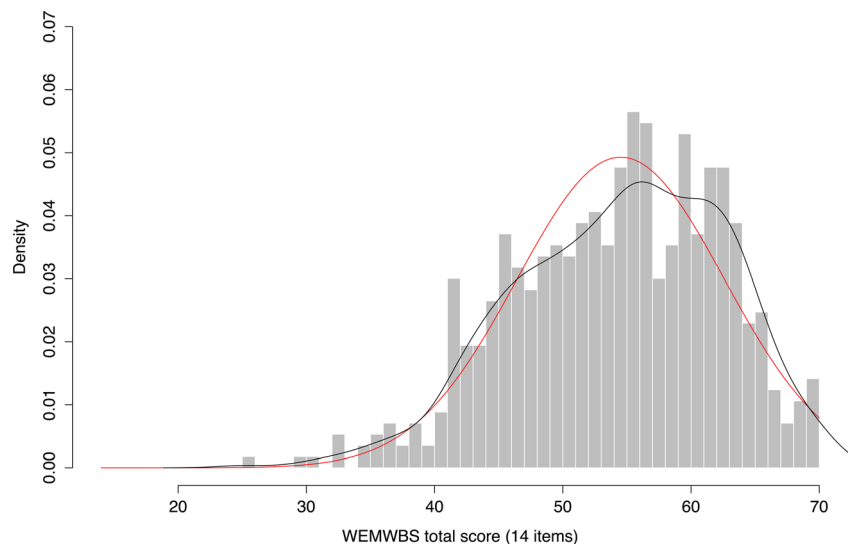
ANOVA: * $p < 0.050$, ** $p < 0.010$, *ns* non-significant (in italics)

empirically justified and recommended. The findings from bi-factor modelling provide important information on how to use and interpret the WEMWBS scores.

Although not hypothesised, three additional measurement error covariances of lower magnitude needed to be specified. This is reasonable because they may arise from similar semantics (i.e. close and intimate relationships of items v9 and v12), be specific for higher age groups (i.e. feeling of being useful and optimistic about the future of items v1 and v2) or be due to translation (e.g. a better translation for “confident” might be “selbstsicher” rather than “zuversichtlich”). Even more significantly, there was no reason to reduce the number of items as has happened in some other studies.

Besides these unique findings regarding construct validity, the German translation showed very similar results when compared to previous validation studies of the WEMWBS. Again, face validity with high response rates, full answer category usage and only slightly skewed answer patterns to the scale items were all replicated in our study. Also reflecting international results, criterion validity was demonstrated with just one quality of life indicator. Yet again, and very similar to previous studies, the German scale version was capable of covering differences in MWB across major socio-demographic groups. In comparison, the total score was slightly skewed but showed neither floor nor ceiling effects as in other studies.

Fig. 1 WEMWBS total score distribution



Conclusions

The present study extends available research about the validity of the WEMWBS by using a bi-factor modelling approach, a psychometric tool that enables the division of specific and common variance. Most of the variance in total WEMWBS scores reflects a single common dimension of MWB. This finding deserves further attention and careful analysis, because it suggests that one could use the composite score to scale individuals on a single dimension of MWB. Interpretation of the WEMWBS total score was psychometrically supported in the present study.

With our study we added to already available scale validations with new information about the measurement quality and the feasibility of applying factorial validity (unidimensionality) to a broad, multi-faceted and heterogeneous indicators measure. Our alternative and advanced modelling strategies provided tremendous additional insights into the scale's performance.

Nevertheless, the present study also has limitations that simultaneously highlight future research priorities: first, since the study sample included participants from a cultural context that might include some specifics of the community-based sample (i.e. supporting members of the Viennese Red Cross of an advanced age living in Vienna), further research on the structural validity of the German WEMWBS in a population-based sample (e.g. in other German-speaking populations or countries) across diverse cultural contexts and wider socio-demographic groups is needed to quantitatively test whether the scale possesses measurement invariance, for example (Steenkamp and Baumgartner 1998). Qualitative scale testing could profit from the application of cognitive interviewing techniques (Willis 2005) as well to confirm item translations and wordings are appropriate, for example regarding item 5 (“I’ve had energy to spare”) or item 10 (“I’ve been feeling confident”). Second, in the present study it was only possible to examine the relation of MWB with one convergent quality-of-life measure (i.e. the EUROHIS-QOL scale). Future research with the German scale version should demonstrate in more depth and detail the external correlations of the WEMWBS, ideally with the general and group factors of the bi-factor model (Chen et al. 2012). Last but not least, the present data are cross-sectional, but in future the longitudinal stability and reliability of the WEMWBS needs further exploration in a structural equation modelling framework, for example (Little 2013).

Above all, the German version of the scale achieved good validity and reliability in our test sample. With this background, the German version should be included in further research activities as a measure for MWB to investigate positive health across (groups of) populations of interest.

Acknowledgements This research would not have been possible without the support of the Viennese Red Cross. We are thankful to Christian Listopad, Astrid Guby and Brigitte Klamert of the Members Support Department (Mitgliederbetreuung) who assisted with the idea to put a special focus on quality of life and well-being in their members’ omnibus survey. Thanks go also to Sarah Stewart-Brown and Frances Taggart from the University of Warwick Medical School for their valuable feedback on an earlier version of this article. We are grateful to the reviewers’ comments and the feedback of Richard Wynne and Keith Chandler in helping us to improve this article. Last but not least, special thanks to all respondents who provided us valuable insights into their lives by participating in this survey.

Compliance with ethical standards

Financial support and sponsorship This study did not receive any funding. It forms part of an internal research project performed by the Viennese Red Cross.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. Informed consent was obtained from all individual participants included in the study.

Conflict of interest The authors Gert Lang and Almut Bachinger declare that they have no conflict of interest.

References

- Azcurra DS (2015) Traducción, adaptación al español y validación de la escala de bienestar mental de Warwick-Edinburgh en una muestra de adultos mayores argentinos. *Acta Colombiana de Psicología* 18:79–93. doi:10.14718/ACP.2015.18.1.8
- Beaton DE, Bombardier C, Guillemin F, Ferraz MB (2000) Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine* 25:3189–3191. doi:10.1097/00007632-200012150-00014
- Bortz J, Döring N (2006) *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler*, 4th edn. Springer, Berlin
- Brähler E, Mühlan H, Albani C, Schmidt S (2007) Teststatistische Prüfung und Normierung der deutschen Versionen des EUROHIS-QOL Lebensqualität-Index und des WHO-5 Wohlbefindens-Index. *Diagnostica* 53:83–96. doi:10.1026/0012-1924.53.2.83
- Brown TA (2006) *Confirmatory factor analysis for applied research*. Guilford Press, New York
- Browne MW, Cudeck R (1993) Alternative ways of assessing model fit. In: Bollen KA, Long JS (eds) *Testing structural equation models*. Sage, Newbury Park, pp 136–162
- Byrne BM (2012) *Structural equation modeling with Mplus. Basic concepts, applications, and programming*. Routledge, New York
- Castellví P et al (2014) The Spanish version of the Warwick-Edinburgh Mental Well-Being Scale (WEMWBS) is valid for use in the general population. *Qual Life Res* 23:857–868. doi:10.1007/s11136-013-0513-7
- Chen FF, West SG, Sousa KH (2006) A comparison of bifactor and second-order models of quality of life. *Multivar Behav Res* 41: 189–225. doi:10.1207/s15327906mbr4102_5
- Chen FF, Hayes A, Carver CS, Laurenceau J-P, Zhang Z (2012) Modeling general and specific variance in multifaceted constructs: a comparison of the bifactor model to other approaches. *J Pers* 80:219–251. doi:10.1111/j.1467-6494.2011.00739.x

- Chen FF, Jing Y, Hayes A, Lee JM (2013) Two concepts or two approaches? A bifactor analysis of psychological and subjective well-being. *J Happiness Stud* 14:1033–1068. doi:10.1007/s10902-012-9367-x
- Clarke A et al. (2011) Warwick-Edinburgh Mental Well-Being Scale (WEMWBS): Validated for teenage school students in England and Scotland. A mixed methods assessment. *BMC Public Health* 11. doi:10.1186/1471-2458-11-487
- Cronbach LJ, Meehl PE (1955) Construct validity in psychological tests. *Psychol Bull* 52:281–302
- de Bruin GP, du Plessis G (2015) Bifactor analysis of the Mental Health Continuum-Short Form (MHCF-SF). *Psychol Rep* 116:438–446. doi:10.2466/03.02.PR0.116k20w6
- Deng N, Guyer R, Ware JE (2015) Energy, fatigue, or both? A bifactor modeling approach to the conceptualization and measurement of vitality. *Qual Life Res* 24:81–93. doi:10.1007/s11136-014-0839-9
- Diener E, Oishi S, Lucas RE (2009) Subjective well-being: the science of happiness and life satisfaction. In: Snyder CR, Lopez SJ (eds) *Oxford handbook of positive psychology*. Oxford University Press, Oxford, pp 63–73. doi:10.1093/oxfordhb/9780195187243.013.0017
- Fromm E (1978–79) Primary and secondary process in waking and in altered states of consciousness. *J Altered States Conscious* 4:125–128
- Gremigni P, Stewart-Brown S (2011) Una misura del benessere mentale: validazione italiana della Warwick-Edinburgh Mental Well-Being Scale (WEMWBS). *Giornale Italiano Di Psicologia* 38:543–563. doi:10.13140/2.1.2425.9203
- Hu L-t, Bentler PM (1999) Cutoff criteria for fit indexes in covariance structure analysis. *Struct Equ Model* 6:1–55. doi:10.1080/10705519909540118
- Hunter SC, Houghton S, Wood LJ (2015) Positive mental well-being in Australian adolescents: evaluating the Warwick-Edinburgh Mental Well-Being Scale. *Aust Educ Dev Psychol* 32:93–104. doi:10.1017/edp.2015.12
- Jovanović V (2015a) A bifactor model of subjective well-being: A re-examination of the structure of subjective well-being. *Personal Individ Differ* 87:49. doi:10.1016/j.paid.2015.07.026
- Jovanović V (2015b) Structural validity of the Mental Health Continuum-Short Form: The bifactor model of emotional, social and psychological well-being. *Personal Individ Differ* 75:154–159. doi:10.1016/j.paid.2014.11.026
- Keldal G (2015) Warwick-Edinburgh Mental İyi Oluş Ölçeği'nin Türkçe Formu: Geçerlik ve güvenilirlik çalışması. *J Happiness Well-Being* 3: 103–115
- Kim S, Han-Yong J, Na K-S, Lee SI, Kim S-G, Lee AR, Cho JT (2014) A validation of the Korean version of the Warwick-Edinburgh Mental Well-Being Scale. *J Korean Neuropsychiatr Assoc* 53:237–245. doi:10.4306/jknpa.2014.53.4.237
- Kline RB (2011) *Principles and practice of structural equation modeling*, 3rd edn. Guilford Press, New York
- Little TD (2013) *Longitudinal structural equation modeling*. Guilford Press, New York
- Lloyd K, Devine P (2012) Psychometric properties of the Warwick-Edinburgh Mental Well-Being Scale (WEMWBS) in Northern Ireland. *J Ment Health* 21:257–263. doi:10.3109/09638237.2012.670883
- Longo Y, Coyne I, Joseph S, Gustavsson P (2016) Support for a general factor of well-being. *Personal Individ Differ*. doi:10.1016/j.paid.2016.03.082
- López MA et al (2013) Adaptation into Spanish of the Warwick-Edinburgh Mental Well-Being Scale (WEMWBS) and preliminary validation in a student sample. *Qual Life Res* 22:1099–1104. doi:10.1007/s11136-012-0238-z
- McDonald RP (1999) Test theory: a unified treatment. Erlbaum, Mahwah
- Mészáros V, Ádam S, Szabó M, Szigeti R, Urbán R (2014) The bifactor model of the Maslach Burnout Inventory-Human Service Survey (MBI-HSS)—An alternative measurement model of burnout. *Stress Health* 30:82–88. doi:10.1002/smi.2481
- Moser CA, Kalton G (1971) *Survey methods in social investigation*. Sage Publications, London
- Muthén LK, Muthén BO (2012) *Mplus*. Statistical analysis with latent variables. User's guide, 7th edn. Muthén & Muthén, Los Angeles
- Power M, Bullinger M, Harper A, WHOQOL-Group (1998) The World Health Organization WHOQOL-100: test of the universality of quality of life in fifteen different cultural groups world-wide. *Health Psychol* 18:495–505
- Reise SP (2012) The rediscovery of bifactor measurement models. *Multivar Behav Res* 47:667–696. doi:10.1080/00273171.2012.715555
- Reise SP, Moore TM, Haviland MG (2010) Bifactor models and rotations: Exploring the extent to which multidimensional data yield univocal scale scores. *J Pers Assess* 92:544–559. doi:10.1080/00223891.2010.496477
- Ryan RM, Deci EL (2001) On happiness and human potentials: a review of research on hedonic and eudaimonic well-being. *Annu Rev Psychol* 52:141–166. doi:10.1146/annurev.psych.52.1.141
- Ryff CD, Keyes CLM (1995) The structure of psychological well-being revisited. *J Pers Soc Psychol* 69:719–727. doi:10.1037/0022-3514.69.4.719
- Santos JJA, Da Costa TA, Guilherme JH, Da Silva WC, Lengler Abentroth LR, Krebs JA, Sotoriva P (2015) Adaptation and cross-cultural validation of the Brazilian version of the Warwick-Edinburgh Mental Well-Being Scale. *Rev Assoc Med Bras* 61: 209–214. doi:10.1590/1806-9282.61.03.209
- Schmidt S, Mühlhan H, Power M (2006) The EUROHIS-QOL 8-item index: psychometric results of a cross-cultural field study. *Eur J Pub Health* 16:420–428. doi:10.1093/eurpub/cki155
- Steenkamp J-BEM, Baumgartner H (1998) Assessing measurement invariance in cross-national consumer research. *J Consum Res* 25:78–90. doi:10.1086/209528
- Stewart-Brown S (2013) The Warwick-Edinburgh Mental Well-Being Scale (WEMWBS): Performance in different cultural and geographical groups. In: Keyes CLM (ed) *Mental well-being. International contributions to the study of positive mental health*. Springer, Dordrecht, pp 133–150
- Stewart-Brown S, Janmohamed K (2008) *Warwick-Edinburgh Mental Well-Being Scale (WEMWBS). User Guide (Version 1)*. Warwick Medical School, NHS Health Scotland, University of Warwick
- Taggart F, Friede T, Weich S, Clarke A, Johnson M, Stewart-Brown S (2013) Cross-cultural evaluation of the Warwick-Edinburgh Mental Well-Being Scale (WEMWBS) - a mixed methods study. *Health and Quality of Life Outcomes* 11. doi:10.1186/1477-7525-11-27
- Tennant R, Fishwick R, Platt S, Joseph S, Stewart-Brown S (2006) Monitoring positive mental health in Scotland: validating the Affectometer 2 scale and developing the Warwick-Edinburgh Mental Well-Being Scale for the UK
- Tennant R et al (2007) The Warwick-Edinburgh Mental Well-Being Scale (WEMWBS): development and UK validation. *Health Qual Life Outcomes* 5:1–13. doi:10.1186/1477-7525-5-63
- Urban D, Mayerl J (2014) *Strukturgleichungsmodellierung. Ein Ratgeber für die Praxis*. VS Verlag für Sozialwissenschaften, Wiesbaden
- Vanhoutte B, Nazroo J (2013) Cognitive, affective and eudemonic well-being in later life: Measurement equivalence over gender and life stage. *Sociol Res Online* 19. doi:10.5153/sro.3241
- Vásquez C, Hervás G (2012) Addressing current challenges in cross-cultural measurement of well-being: the Pemberton Happiness Index. In: Knoop HH, Delle Fave A (eds) *Well-being and cultures: perspectives from positive psychology*. Springer, p 31–50
- Waqas A et al (2015) Measuring the well-being of health care professionals in the Punjab: a psychometric evaluation of the Warwick-Edinburgh Mental Well-being Scale in a Pakistani population. *PeerJ* 3. doi:10.7717/peerj.1264

- Warwick Medical School (2016) Warwick-Edinburgh Mental Well-Being Scale (WEMWBS). <http://www2.warwick.ac.uk/fac/med/research/platform/wemwbs/>. Accessed 28. Mar. 2016
- Waterman AS (1993) Two conceptions of happiness: Contrasts of personal expressiveness (eudaimonia) and hedonic enjoyment. *J Pers Soc Psychol* 64:678–691. doi:10.1037/0022-3514.64.4.678
- wien.at (2014) Population and surface of Vienna's municipal districts 2014. Vienna City Administration. <http://www.wien.gv.at/english/administration/statistics/population-district.html>. Accessed 9. Aug. 2014
- Willis GB (2005) Cognitive interviewing. A tool for improving questionnaire design. Sage, London
- Zinbarg RE, Revelle W, Yovel I, Li W (2005) Cronbach's α , Revelle's β , and McDonald's ω H: their relations with each other and two alternative conceptualizations of reliability. *Psychometrika* 70:123–133. doi:10.1007/s11336-003-0974-7