

DEVELOPMENT OF TWO EQUIVALENT SHORT FORMS OF THE PSYCHOLOGICAL GENERAL WELL-BEING INDEX: PGWBI-A AND PGWBI-B

SILVIA TESTA
CRISTINA CIVILOTTI
GIULIA DI FINI
UNIVERSITY OF TORINO

CRISTINA ROSSETTO
CENTER FOR THERAPEUTIC STUDIES OF EMOTIONAL AND SEXUAL DISEASES
(DAS CENTER) OF GENOVA

VIERI BONCINELLI
UNIVERSITY OF PISA

FABIO VEGLIA
UNIVERSITY OF TORINO

Psychological General Well-Being Index (PGWBI) is a generic health-related quality of life instrument that produces subscale scores about six domains of well-being and a global score. A literature review showed that the factor validity was good for the global score but was poor for the subscale scores. The aim of the study was to develop two short forms of the instrument that produce equivalent global well-being scores. Data gathered from two samples of healthy people coming from the North, the Center, and the South of Italy were employed for the development ($N = 428$) and for the assessment of measurement invariance ($N = 220$) of the two forms. Confirmatory factor analysis (CFA) was used to derive the two forms and to test their invariance. The two forms (PGWBI-A and PGWBI-B) showed items' partial measurement invariance and their global scores showed good internal coherence.

Key words: Well-being; Short form; Assessment; PGWBI; Quality of life.

Correspondence concerning this article should be addressed to Giulia Di Fini, Department of Psychology, University of Torino, Via Po 14, 10123 Torino, Italy. Email: giulia.difini@unito.it

Quality of life (QoL) has been defined by the World Health Organization (WHO) as “individual’s perception of their position in life in the context of the culture and value systems in which they live and in relation to their goals, expectations, standards and concerns” (WHOQOL Group, 1995, p. 1405). In the last two decades of the twentieth century, health research introduced the concept of subjective quality of life in order to assess the consequences of chronic diseases from a medical point of view and to get an evaluation of the efficacy of medical treatments (Bowling, 1995). The increase in life expectancy due to medical progress, along with the preva-

lence of chronic diseases and inadequate medical measures for the assessment of effects of treatment, are just some of the factors that increased the interest in evaluating the quality of life in medical settings (WHOQOL Group, 1994).

From 1980 onwards, many instruments have been developed for the assessment of functioning in everyday life. The last two decades have seen an ongoing debate in QoL research, especially with regard to its assessment, which has led to agreement on some principles (Breslin, 1991; Ready, 2002): QoL should be measured in a comprehensive manner, covering a wide range of aspects; it is important to incorporate a subjective measure of self-report; instruments for QoL assessment should be culturally sensitive, feasible, and rapid in their administration.

Among the instruments assessing health-related quality of life (HRQoL), some are categorized as generic, while others as condition-specific (Coons, Rao, Keininger, & Hays, 2000; Guyatt, Feeny, & Patrick, 1993). The advantage of generic measures consists of the possible use among different populations, with a wide range of medical conditions. The Psychological General Well-Being Index (PGWBI) can be considered as one of the first generic tools for the evaluation of QoL. It was designed by Harold Dupuy, psychologist for the National Center for Health Statistics, in the late Sixties, in order to meet the U.S. government's need for a simple but valid and reliable measure of the American population's emotional and physiological distress and self-perception of affective states. The first version consisted of 68 items. Thereafter the number of items was reduced to 18 to ensure a wider acceptability of the questionnaire. The final validated version, named PGWB Index, was born from the collaboration between Dupuy and Ware and contains 22 selected items with response categories normalized according to a 6-point Likert scale, with values between 0 and 5 (Dupuy, 1984). Six subscale scores covering the dimensions of Anxiety, Depression, Positive well-being, Vitality, General health, and Self-control, and a global score (Index) can be calculated.

In the early Nineties, the PGWBI was introduced in Europe and was adapted into many languages under the coordination of the MAPI Research Institute. The Italian validation of the PGWBI was carried out in 2000 within the project MIOS, which aimed at understanding the value of some measures based on the patient's subjectivity (Grossi, Mosconi, Groth, Niero, & Apolone, 2002). Later, Grossi and colleagues (2006), using a stepwise regression approach, validated the abbreviated version of the PGWBI (PGWB-S) in Italian samples, in order to increase the acceptability among the questionnaire respondents.

One of the advantages of PGWBI is its ability to measure psychological states that are both lower and higher than the normal levels and which are normally linked to the simple lack of distress. It provides an index that measures the self-representation of well-being, or emotional and affective distress, which, thanks to its sensitivity to changes, has shown to be useful both during measurements on normal populations and during health examination programs or after medical treatments.

The PGWBI has been used in several studies and with a broad range of patient groups, such as patients with cardiovascular disease, gastrointestinal symptoms, sexual dysfunction, cancer, and so forth (Casellas, Vivancos, Sampedro, & Malagelada, 2005; Enck, Dubois, & Marquis, 1999; Herlitz et al., 1998; Hultling et al., 2000; Saner, Borner Rodriguez, Kummer-Bangerter, Schuppel, & von Planta, 2002; Serpentine et al., 2011). However, from a literature review it emerges that there is still a lack of general agreement on the factor structure of the PGWBI. In

particular, the exploratory factor analysis (EFA) did not seem to reproduce the six factors, recommended for clinical use that were mentioned in the operating instructions of the instrument. (Dupuy, 1984; Revicki, Shakespeare, & Kind, 1996). For example, Wool, Cerutti, Marquis, Cialdella, and Hervie (2000), after administering the PGWBI to a sample of 155 Italian menopausal women, using principal component analysis (PCA), identified a structure with a single general factor and then, after rotation, three factors. On the other hand, factor analysis performed by Gaston and Vogl (2005), using the British version of the PGWBI — General Well-Being Index (GWBI) administered to a sample of students, yielded a three-factor solution. The three-factor solution was not the same across different studies (Gaston & Vogl, 2005; Wool et al., 2000), however, a single general factor of well-being seems to be predominant. PGWBI was implicitly assumed to assess a unidimensional construct of psychological well-being in Grossi et al.'s (2006) development of the short form PGWB-S. More recently, Lundgren-Nilsson, Jonsdottir, Ahlborg, and Tennant (2013) assessed construct validity by means of Rasch models and factor analysis. They obtained that: EFA yielded two factors (distinguishing positive and negative domains); confirmatory factor analysis (CFA) did not support the unidimensional model, but the fit of the unidimensional model was good when items of each domain were parceled; Rasch analysis performed by using testlets produced a satisfactory unidimensional solution.

All these studies show that it is difficult to obtain each of the six theoretical domains empirically, whereas the global PGWBI score is a useful measure of personal well-being. For this reason, a shortening of the instrument would be valuable, since 22 items are too many to measure a single dimension. The interest in the development of a short-form of psychological well-being instruments has grown in the last decade (Goetz et al., 2013). It is becoming more and more evident that assessment instruments should be valid and reliable but also efficient, that is, acquiring the data should not take more time than strictly necessary in order to save time and money and to maintain the compliance of the examinee. Item reduction could be performed by several statistical approaches some of which have been recognized as inappropriate since they could result in selection of a wrong item. Cronbach's alpha and other maximizing internal consistency criteria, for example, lead to the selection of redundant items and stepwise regression is heavily affected by the part-whole correlation effect and may erroneously eliminate items with the lowest measuring error (Coste, Guillemin, Pouchot, & Fermanian, 1997; Goetz et al., 2013; Stanton, Sinar, Balzer, & Smith, 2002). In the development of a short form, it is also important to preserve the content coverage of the instrument (Smith, McCarthy, & Anderson, 2000).

The preliminary aim of this study was to assess what model better fitted the data between the theoretical model with six first-order factors and one second-order factor and the unidimensional one, in order to determine the starting model for the shortening of the instrument. The main objective of the study was the formulation of two equivalent short forms of the PGWBI, that is, two short psychological well-being scales whose scores are characterized by equivalent precision and by equivalent mean value. The existence of two equivalent forms could be very useful from both a clinical and a research standpoint because it could allow for the administration of the PGWBI in repeated measure design without causing addiction in the respondent, regardless of the time period between the two administrations.

METHOD

Participants and Procedure

The development of the two short forms was done in two steps. The first step was devoted to the development of the two forms on a sample of 428 participants. In the second step, the measurement invariance of the two short forms was assessed in an independent sample of 220 participants.

The study presented here was derived from a larger investigation promoted by the Italian Federation of Scientific Sexology (FISS) with the goal of evaluating general well-being in relation to the subjective perception of sexuality in the general population. Participants were recruited from several cities in a convenience sample, from the North, the Center, and the South of Italy, using affiliated FISS centers as bases for recruitment. The socioanagraphic characteristics of the two samples are reported in Table 1.

TABLE 1
Sample characteristics

Variables		Step 1		Step 2	
		Development sample (<i>N</i> = 428)		Cross-validation sample (<i>N</i> = 220)	
		<i>n</i>	%	<i>n</i>	%
Gender	Male	196	45.8	75	34.1
	Female	214	50.0	143	65.0
	Information missing	18	4.2	2	0.9
Age	20-30	187	43.7	78	35.4
	31-40	136	31.8	49	22.3
	41-50	105	24.5	58	26.4
	Over 50	–	–	32	14.5
	Information missing	–	–	3	1.4
Education	Elementary school	–	–	4	1.8
	Junior high school	52	12.1	32	14.5
	High school	197	46.0	79	35.9
	College/University	179	41.8	104	47.3
	Information missing	–	–	1	0.5
Residence	North	317	74.1	131	59.5
	Center	52	12.1	34	15.5
	South and Islands	40	9.3	48	21.8
	Information missing	19	4.4	7	3.2

In both steps, participants completed a self-administrated questionnaire composed of a socioanagraphic section, the PGWBI instrument, and scales evaluating sexual satisfaction and practices and quality of marital or dyadic relationships. In this context, only the PGWBI data will be analyzed.

Instruments

In Step 1, the Italian version of PGWBI was used (Grossi et al., 2002). It was composed of 22 items that give rise to six subscale scores, covering the dimensions of Anxiety (five items), Depression (three items), Positive well-being (four items), Vitality (four items), General health (three items), and Self-control (three items), and a global score (22 items). Item responses were scored on a 6-point scale (10 items were in reverse form). Cronbach's alpha for the 22-item global score was .94.

In Step 2, the two six-item short forms, PGWBI-A and PGWBI-B, obtained in the previous step were employed. Items of both forms were scored on a 6-point scale. Cronbach's alphas were .81 and .82, respectively.

STATISTICAL ANALYSIS

Step 1a: PGWBI Dimensionality Investigation

Statistical analysis of Step 1 involved 388 of the 428 participants who completed the questionnaire: we included in the analysis the participants with valid responses on all the PGWBI items. Two confirmatory factor models (Lisrel 8.72; Jöreskog & Sörbom, 1996) were estimated: the original theoretical model with six first-order factors and one second-order factor, and the unidimensional model. Since data violated the normality assumption, Mardia's test ($2, 388 = 2803.01, p < .001$), a robust maximum likelihood (ML) estimator applied to the variance-covariance matrix (Satorra & Bentler, 1994) was used.¹

Several criteria were employed to evaluate the two models: the root mean square error of approximation (RMSEA), the comparative fit index (CFI), and the standardized root mean square residual (SRMR). The following rule of thumb was used to judge model goodness of fit as satisfactory: RMSEA < .08; CFI > .95; SRMR < .08 (Browne & Cudeck, 1993; Hu & Bentler, 1995, 1996). To compare the goodness of fit of the two models, the Satorra and Bentler (2001) scaled difference chi-square test was employed.

Step 1b: Development of the Two Short Forms

Classical Test Theory (CTT) approach was followed in the development of the short forms. According to CTT (Lord & Novick, 1968), two items are defined as "strictly parallel forms" if they measure the same construct with equal discrimination (i.e., the latent factor has the same impact on them, they have equal loadings), equal precision (i.e., equal error variances and, therefore, equal reliability), and equal difficulty (difficulty is quantified in relation to the proportion of people who choose the different response option to an item, and it is measured through item's mean). Items are "essentially parallel measures" if equality on difficulty is relaxed and "tau-equivalent forms" (strictly or essentially) if equality of error variances is relaxed. The goal was to select two strictly parallel (or at least essentially tau-equivalent) items for each of the six content domains of which the PGWBI total score is composed, and then to create two short forms of PGWBI, each composed of six items selected from the six couples of items.

A CFA model was specified, imposing equality constraints on loadings, error variances and means for subsets of items belonging to the same theoretical domain (Anxiety, Depression, Positive well-being, Vitality, General health, Self-control). By examining modification indices, the constraint that mostly contributed to model unfitting was detected and the model was re-estimated without it. This sequence was replicated until the global fit of the model was acceptable, ensuring however that there were at least two items per each content domain. In the end, this procedure yielded two items for each content domain, which were split into two PGWBI short forms of six items each, whose performance was evaluated by two unidimensional CFA models.

The equivalence of the two short form scores was evaluated by means of paired *t*-tests, Cohen's *d* (Lakens, 2013), and intraclass correlation coefficient (ICC). The internal coherence was assessed by Cronbach's alpha. The Spearman-Brown prophecy formula was used to compare internal coherence of the short form scores with that of the 22-item general score.

Step 2: Assessment of Measurement Invariance of the Two Short Forms

Statistical analysis was performed on data from 211 of the 220 participants involved in Step 2: nine participants were excluded because of one or more missing data regarding the two PGWBI short forms. The factor invariance of the two forms was assessed by following the procedure recommended by Vandenberg and Lance (2000). Specifically, the hierarchical sequence of invariance tests suggested for longitudinal data was used; the present situation in which two equivalent forms were evaluated on the same sample in the same occasion was considered as equivalent to the situation in which the same form was administered to the same sample in two different occasions.

First, an overall test of invariance was performed, constraining covariance matrices and means to be equal across the two forms (Model 0). If this test had not been passed, a set of sequential tests would have been performed in order to comprehend which aspects were responsible for the measurement inequivalence. The sequence of tests was the following: configural invariance, that is, the same pattern of loadings across forms (Model 1); metric invariance, that is, the same values for the loadings across forms (Model 2); scalar invariance, that is, equality of item intercepts across forms (Model 3); uniqueness invariance, that is, equality of error variances across forms (Model 4); factor variance equality across forms (Model 5), and equality of factor means across the two forms (Model 6). To compare subsequent nested models, the abovementioned Satorra and Bentler scaled difference chi-square test was applied. As in Step 1, the equivalence of the two short form scores was evaluated by means of paired *t*-tests, Cohen's *d*, and ICC. Lisrel 8.72 and SPSS22 were employed for the analysis.

RESULTS

Step 1a: PGWBI Dimensionality Investigation

The unidimensional model and the hierarchical model with six first-order factors and one second-order factor were estimated by confirmatory factor analysis. Both models adequately fitted the 22-item covariance matrix. As shown in Table 2, in both cases, the values of RMSEA, CFI, and SRMR were within acceptability ranges.²

TABLE 2
Step 1: CFA goodness of fit measures (22-items PGWBI)

Model	Satorra-Bentler χ^2	df	p	RMSEA [90% CI]	CFI	SRMR
Hierarchical model	416.83	203	< .01	.05 [.05, .06]	.98	.05
Unidimensional model	618.42	209	< .01	.07 [.07, .09]	.96	.06

Note. CFA = confirmatory factor analysis; RMSEA [CI] = root mean square error of approximation [confidence interval]; CFI = comparative fit index; SRMR = standardized root mean square residual.

The Satorra and Bentler scaled difference chi-squared test used to compare the fit of the two models yielded a statistically significant result, scaled difference test (6) = 255.72, $p < 0.01$, meaning that the less constrained model (the hierarchical one) better fitted the data. In light of this last result, the hierarchical model was used in the development of the two forms.

Step 1b: Development of the Two Short Forms

The procedure described above in the Statistical Analysis section was applied to the hierarchical model, with the aim of selecting two items per content domain that would respect the CTT condition of parallelism or tau-equivalence. Equality constrains on loadings, error variances or means that mostly contributed to the unfitting of the model were relaxed one at a time and the model was re-estimated. The starting model with constrains on all of the 22 items was far from being satisfactory, since RMSEA, CFI, and SRMR were very distant from their threshold values: RMSEA = .18, 90% CI [.17, .18]; CFI = .72; SRMR = .13. The goodness of fit of the final model composed of 12 items was acceptable: RMSEA = .07, 90% CI [.05, .08]; CFI = .96; SRMR = .07. Five couples of items were strictly parallel, while the two items of Self-control factor (Item Q4r and Item Q18)³ resulted to be essentially parallels since the equality of means had to be removed.

The 12 items were split into two sets of six, called PGWBI-A and PGWBI-B. The order in the questionnaire was used as the general rule to assign items to PGWBI-A or PGWBI-B: for each couple, the item that came first in the questionnaire was put in PGWBI-A, with one exception (Items Q20-Q9r) due to the necessity to balance the number of reverse items in the two forms. The two short forms are reported in Appendix A (English version) and Appendix B (Italian version).

The performance of the CFA unidimensional model applied separately to the two forms was very good. As shown in Figure 1, all the loadings were statistically significant and high, ranging from .51 to .80, and fit index values were excellent.

The summary statistics for the items composing the two forms and of the two global scores are reported in Table 3. Three couples of items showed a statistically significant means difference, but only the difference between Q4r and Q18 means deserved attention, since Cohen's d was of medium size ($d = .50$). The difference between the means scores of the two forms (3.65 and 3.60, respectively for PGWBI-A and PGWBI-B) was statistically significant, however, its effect size was negligible (Cohen's $d = .17$).

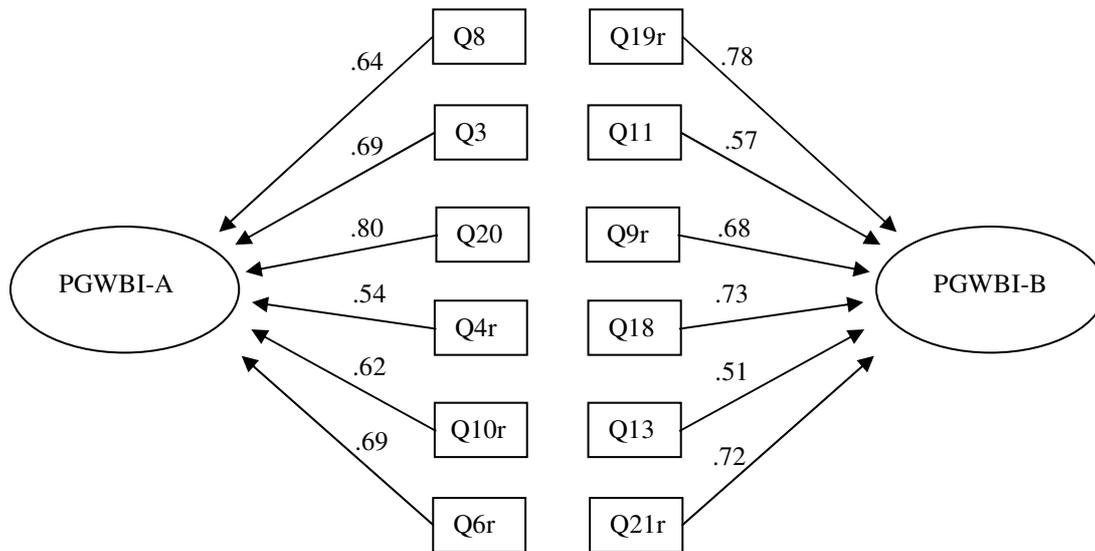


FIGURE 1

Step 1: CFA path diagrams of the unidimensional solutions of the two PGWBI short forms (standardized loadings, all statistically significant, $p < .01$).
PGWBI-A: SB $\chi^2(9) = 10.82, p = .80$; RMSEA = .00, 90% CI [.00, .04]; CFI = 1.00; RMRS = .02.
PGWBI-B: SB $\chi^2(9) = 20.98, p = .09$; RMSEA = .04, 90% CI [.00, .08]; CFI = .99; RMRS = .03.
r = reverse item.

TABLE 3
Step 1: Means and standard deviations of the two PGWBI short forms

	PGWBI-A		PGWBI-B		Paired means Cohen's <i>d</i>	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
Q8	3.11	1.15	Q19r	3.14	1.06	-.03
Q3	4.36 ^a	0.83	Q11	4.50 ^a	1.00	-.20
Q20	3.08 ^b	1.08	Q9r	2.95 ^b	0.97	.15
Q4r	3.81 ^c	1.13	Q18	3.31 ^c	1.06	.50
Q10r	4.05	0.88	Q13	4.05	1.00	.00
Q6r	3.52	0.93	Q21r	3.64	1.03	-.11
Mean score	3.65 ^d	0.73	Mean score	3.60 ^d	0.75	.17

Note. ^a $t(387) = -3.66, p < .01$; ^b $t(387) = -3.00, p < .01$; ^c $t(387) = 9.73, p < .01$; ^d $t(387) = -3.22, p < .01$.
r = reverse item.

The ICC between PGWBI-A and PGWBI-B was .94. Cronbach's alphas were .82 and .83, very close to .81 obtained applying the Spearman-Brown prophetic formula to the 22-item internal coherence value of .94.

Step 2: Assessment of Measurement Invariance of the Two Short Forms

In Step 2, the equivalence of the two short forms was assessed in an independent sample, employing a hierarchical set of invariance tests (Tables 4a and 4b). As showed in Table 4a, Model 0 that imposes a full invariance across the two forms provided an unsatisfactory fit to the data: CFI and SRMR values were good, but RMSEA value (.10) was outside of the recommended range (.00-.08). This result suggested the presence of some form of inequivalence and Models from 1 to 6, whose results are reported in Table 4a, were estimated to identify the source of the inequivalence. Model 1 (configural invariance) provided excellent fit to the data by every fit index. Also the fit of Model 2 (metric invariance) was acceptable, but the statistically significant $\Delta\chi^2$ test suggested that some loading constrains worsened model fit. Model 3, that imposes intercept equality constrains (in addition to loading constrains), showed acceptable fit values, but as for Model 2, the $\Delta\chi^2$ test was statistically significant, meaning that some intercept constrains did not fit the data. The same happened when uniqueness invariance was assessed (Model 4). Finally, Models 5 and 6, testing the equality of factor variances and factor means demonstrated good fit according to the $\Delta\chi^2$ test and all the fit indexes.

Summarizing, the six models provided acceptable fit values, but $\Delta\chi^2$ test evidenced some sources of inequivalence related to loadings, intercepts, and uniqueness. A close examination of Model 2 showed that the two forms were partially metric invariant once the equality constrain between Q20 and Q9r items had been relaxed (Table 4b). Furthermore, to obtain a statistically non-significant $\Delta\chi^2$ test for scalar invariance it was necessary to relax intercept equality constrains from the following three couples of items Q20-Q9r, Q4r-Q18, and Q3-Q11. Finally, partial uniqueness invariance was obtained relaxing the error variance constrain for items Q8 and Q19r.

The ICC between PGWBI-A and PGWBI-B was .92. Regarding means' differences (Table 5), three couples of items showed a statistically significant means' difference, but only the difference between the means of Q3 and Q11 showed an appreciable effect size ($d = -.34$), whereas the difference between the means' scores of the two forms was not statistically significant.

TABLE 4a
Step 2: Assessment of measurement invariance across PGWBI-A and PGWBI-B

Model	Satorra-Bentler χ^2	df	RMSEA [90% CI]	CFI	SRMR	$\Delta\chi^2$	Δdf
0 Invariant covariance matrices	79.29	27	.10 [.07, .12]	.97	.07	–	–
1 Configural invariance	60.30	47	.04 [.00, .06]	.99	.04	–	–
2 Metric invariance	82.89	52	.05 [.03, .07]	.98	.08	32.83*	5
3 Scalar invariance	117.08	57	.07 [.05, .09]	.97	.08	80.34*	5
4 Uniqueness invariance	138.48	63	.08 [.06, .09]	.96	.08	18.39*	6
5 Factor variances' invariance	139.16	64	.08 [.06, .09]	.96	.08	0.20	1
6 Factor means' invariance	142.79	65	.08 [.06, .09]	.96	.08	4.35	1

Note. RMSEA [CI] = root mean square error of approximation [confidence interval]; CFI = comparative fit index; SRMR = standardized root mean square residual.

* $p < .01$.

TABLE 4b
Step 2: Tests of partial invariance across PGWBI-A and PGWBI-B

Model	Satorra-Bentler χ^2	df	RMSEA [90% CI]	CFI	SRMR	$\Delta\chi^2$	Δf
1 Configural invariance	60.30	47	.04 [.00, .06]	.99	.04	–	–
2a Partial metric invariance (a)	67.07	51	.04 [.00, .06]	.99	.05	8.09	4
3a Partial scalar invariance (b)	72.33	53	.04 [.01, .06]	.99	.05	6.47	2
4a Partial uniqueness invariance (c)	83.81	58	.05 [.02, .07]	.99	.06	9.66	5
5 Factor variances' invariance	86.42	59	.05 [.02, .07]	.98	.07	2.51	1
6 Factor means' invariance	89.30	60	.05 [.03, .07]	.98	.07	3.83	1

Note. RMSEA [CI] = root mean square error of approximation [confidence interval]; CFI = comparative fit index; SRMR = standardized root mean square residual. (a) loadings' equality constrain relaxed from the couple Q20-Q9r; (b) intercepts' equality constrain relaxed from the couples Q3-Q11, Q20-Q9r, Q4r-Q18; (c) error variances' equality constrain relaxed from the couple Q8-Q19r.

TABLE 5
Step 2: Means and standard deviations of the two PGWBI short forms

PGWBI-A			PGWBI-B			Paired means Cohen's <i>d</i>
	<i>M</i>	<i>SD</i>		<i>M</i>	<i>SD</i>	
Q8	2.99 ^a	1.11	Q19r	3.14 ^a	0.99	-.19
Q3	4.31 ^b	0.84	Q11	4.58 ^b	0.88	-.34
Q20	3.13	1.13	Q9r	2.99	1.04	.14
Q4r	3.74 ^c	1.18	Q18	3.51 ^c	1.16	.19
Q10r	4.01	0.85	Q13	3.92	1.04	.09
Q6r	3.49	0.92	Q21r	3.53	1.00	-.04
Mean score	3.61	0.72	Mean score	3.61	0.75	.00

Note. ^a $t(210) = -2.71, p < .01$; ^b $t(210) = -4.92, p < .01$; ^c $t(210) = 2.74, p < .01$.

r = reverse item.

DISCUSSION

PGWBI is one of the first generic instruments proposed to assess health-related QoL and has been used in several studies involving both healthy people and people suffering from a broad range of diseases. It is composed of 22 items, covering six content domains: Anxiety, Depression, Positive well-being, Vitality, General health, and Self-control. Albeit it allows the computation of six subscale scores corresponding to the content domains, the factor validity of the six factors' scores was poor; a literature review showed that when exploratory factor analysis was performed, one, two, or three factors instead of six emerged (Gaston & Vogl, 2005; Lundgren-Nilsson et al., 2013; Wool et al., 2000). At the same time, the 22 items produced a global score that proved to

be a valid measure of personal well-being, as resulted, for example, from the Rasch analysis by Lundgren-Nilsson and colleagues.

Since 22 items are a lot to assess a single general measure of well-being, the aim of this study was the development of two equivalent PGWBI short forms, each composed of six items, one item per content domain. The selected 6+6 items showed acceptable psychometric properties, albeit some misfit was observed in both steps of the analysis. In Step 1, the items referring to Self-control domain resulted to be essentially parallel forms since their intercepts were different and their means' difference was also appreciable in terms of effect size. In Step 2 items showed partial measurement invariance. Specifically, the loadings of the items belonging to the Positive well-being domain (Q20 and Q9r) were not equal; the precision (error variances) of the items belonging to the Anxiety domain (Q8 and Q19r) were not equal and the intercepts of the items belonging to the Positive well-being domain, Depression, and Self-control (Q20-Q9r, Q3-Q11, and Q4r-Q18, respectively) were not equal.

These misfits did not seem to threaten the usability of the two forms as equivalent instruments and this for several reasons. In terms of goodness of fit indexes, all the strictly invariance models of Step 2 showed acceptable values and the evidence of partial invariance emerged only according to the inferential χ^2 difference test. Problematic items were not the same across the two steps: in Step 1 items referring to Self-control were problematic in terms of means' difference, whereas in Step 2 items referring to Depression were problematic in terms of means' difference, and in Step 1 none of the items was problematic in terms of error variances nor loadings' invariance.

The performance of the two global scores was also acceptable. The congruence between the two scores was very good, both in Step 1 (ICC = .94) and in Step 2 (ICC = .92), and the means scores' difference was small in size: it was statistically significant, but negligible in terms of effect size, in Step 1 and not statistically significant in Step 2.

In summary, the two scales may be considered as equivalent forms that could be used in repeated measures studies. The two short forms are composed of six items like the form developed by Grossi and colleagues (2006). They differ from the latter for the principle that guides the item selection. Grossi and colleagues selected the six items that reproduced the maximum amount of the 22-item score variance, covering five out of the six content domains (psychological distress due to physical health was left out). In the development of the two forms, we chose to cover all of the six domains since we believe the content richness of the PGWBI scale to be one of its strengths.

With regard to study implications, the short version lessens the time of administration and increases the acceptability among the questionnaire's respondents. Furthermore, the development of two equivalent measures can be useful in situations where the assessment must be repeated over time. There are some limitations, which should be mentioned. The two forms were administered to the same samples, and further studies should investigate their performance when separately administered to independent and random selected samples. Moreover, longitudinal measurement invariance and test-retest reliability should be assessed to further prove the applicability in repeated measures studies. Additional studies should extend the validation of the instruments to samples of unhealthy people and to other countries.

NOTES

1. The original theoretical model with six first-order factors and one second-order factor was also estimated by full information maximum likelihood (FIML) method, in order to assess the potential bias due to missing deletion.
2. In reference to the supplementary analysis explained in Note 1, the absolute difference between FIML and ML point estimates for first- and second-order loadings and intercepts were marginal ($M = .01$, $SD = 0.01$; min = 0.00; max = 0.03).
3. Here and in the following text, the suffix “r” after the item label indicates reverse-coded data.

ACKNOWLEDGEMENTS

The authors wish to thank the Italian Federation for Scientific Sexology (FISS), especially prof. Vieri Boncinelli and dr. Cristina Rossetto for the logistic and organizational support.

FUNDING

The study was partially supported by Edo e Elvo Tempia Foundation, Biella, Italy.

REFERENCES

- Bowling, A. (1995). What things are important in people's lives? A survey of the public's judgments to inform scales of health related QoL. *Social Science and Medicine*, *41*, 1447-1462. doi:10.1016/0277-9536(95)00113-L
- Breslin, S. (1991). Quality of life: How is it measured and defined? *Urologia Internationalis*, *46*, 246-251. doi:10.1159/000282146
- Browne, M. W., & Cudeck, R. (1993). Alternatives ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 132-162). Thousand Oaks, CA: Sage.
- Casellas, F., Vivancos, J. L., Sampedro, M., & Malagelada J. R. (2005). Relevance of the phenotypic characteristics of Crohn's disease in patient perception of health-related quality of life. *The American Journal of Gastroenterology*, *100*, 2737-2742. doi:10.1111/j.1572-0241.2005.00360.x
- Coons, S. J., Rao, S., Keininger, D. L., & Hays, R. D. (2000). A comparative review of generic quality-of-life instruments. *Pharmacoeconomics*, *17*, 13-35. doi:10.2165/00019053-200017010-00002
- Coste, J., Guillemin, F., Pouchot, J., & Fermanian, J. (1997). Methodological approaches to shortening composite measurement scales. *Journal of Clinical Epidemiology*, *50*, 247-252. doi:10.1016/S0895-4356(96)00363-0
- Dupuy, H. J. (1984). The Psychological General Well-Being (PGWB) index. In N. K. Wenger, M. E. Mattson, C. D. Furberg, ... J. Elinson (Eds.), *Assessment of quality of life in clinical trials of cardiovascular therapies* (pp. 170-183). Retrieved from https://www.openlibrary.org/works/OL9259175W/Assessment_of_Quality_of_Life_in_Clinical_Trials_of_Cardiovascular_Therapies
- Enck, P., Dubois, D., & Marquis, P. (1999). Quality of life in patients with upper gastrointestinal symptoms: Results from the Domestic/International Gastroenterology Surveillance Study (DIGEST). *Scandinavian Journal of Gastroenterology*, *231*, 48-54. doi:10.1080/003655299750025264
- Gaston, J. E., & Vogl, L. (2005). Psychometric properties of the General Well-Being Index. *Quality of Life Research*, *14*, 71-75. doi:10.1007/s11136-004-0793-z
- Goetz, C., Coste, J., Lemetayer, F., Rat, A. C., Montel, S., Recchia, S., ... Guillemin, F. (2013). Item reduction based on rigorous methodological guidelines is necessary to maintain validity when shortening composite measurement scales. *Journal of Clinical Epidemiology*, *66*, 710-718. doi:10.1016/j.jclinepi.2012.12.015

- Grossi, E., Groth, E., Mosconi, P., Cerutti, R., Pace, F., Compare, A., & Apolone, G. (2006). Development and validation of the short version of the Psychological General Well-Being Index (PGWB-S). *Health and Quality of Life Outcomes*, 4, 88. doi:10.1186/1477-7525-4-88
- Grossi, E., Mosconi, P., Groth, N., Niero, M., & Apolone, G. (2002). *Questionario Psychological General Well-Being Index. Versione italiana* [The Psychological General Well Being Questionnaire. Questionnaire for the evaluation of the general psychological well-being]. Milano, Italy: Istituto di Ricerche Farmacologiche "Mario Negri".
- Guyatt, G. H., Feeny, D. H., & Patrick, D. L. (1993). Measuring health-related quality of life. *Annals of Internal Medicine*, 118, 622-629. doi:10.7326/0003-4819-118-8-199304150-00009
- Herlitz, J., Haglid, M., Wiklund, I., Caidahl, K., Karlson, B. W., Sjolund, H., & Karlsson, T. (1998). Improvement in quality of life during five years after coronary artery bypass grafting. *Coronary Artery Disease*, 9, 519-526.
- Hu, L., & Bentler, P. M. (1995). Evaluating model fit. In R. H. Hoyle (Ed.), *Structural equation modelling: Concepts, issues, and applications* (pp. 76-99). Thousand Oaks, CA: Sage.
- Hu, L., & Bentler, P. M. (1996). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1-55. doi:10.1080/10705519909540118
- Hultling, C., Giuliano, F., Quirk, F., Pena, B., Mishra, A., & Smith, M. D. (2000). Quality of life in patients with spinal cord injury receiving Viagra (sildenafil citrate) for the treatment of erectile dysfunction. *Spinal Cord*, 38, 363-370.
- Jöreskog, K. G., & Sörbom, D. (1996). *LISREL 8: User's reference guide*. Chicago, IL: Scientific Software International.
- Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for *t*-tests and ANOVAs. *Frontiers in Psychology*, 4, 863. doi:10.3389/fpsyg.2013.00863
- Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Reading, Mass: Addison-Wesley.
- Lundgren-Nilsson, A., Jonsdottir, I. H., Ahlberg, G., & Tennant, A. (2013). Construct validity of the Psychological General Well-Being Index (PGWBI) in a sample of patients undergoing treatment for stress-related exhaustion: A Rasch analysis. *Health and Quality of Life Outcomes*, 11, 1. doi:10.1186/1477-7525-11-2
- Ready, R. E. (2002). Quality of life in dementia. *Medicine and Health*, 85, 216-218.
- Revicki, D. A., Shakespeare, A., & Kind, P. (1996). Preferences for schizophrenia-related health states: A comparison of patients, caregivers and psychiatrists. *International Clinical Psychopharmacology*, 11, 101-108.
- Saner, H., Borner Rodriguez, E., Kummer-Bangerter, A., Schuppel, R., & von Planta, M. (2002). Quality of life in long-term survivors of out-of-hospital cardiac arrest. *Resuscitation*, 53, 7-13. doi:10.1016/S0300-9572(02)00002-3
- Satorra, A., & Bentler, P. M. (1994). Corrections to test statistics and standard errors in covariance structure analysis. In A. Von Eye & C. C. Clogg (Eds.), *Latent variable analysis: Applications for developmental research*. Thousand Oaks, CA: Sage.
- Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*, 66, 507-514. doi:10.1007/BF02296192
- Serpentini, S., Del Bianco, P., Alducci, E., Toppan, P., Ferretti, F., Folin, M., ... Pucciarelli, S. (2011). Psychological well-being outcomes in disease-free survivors of mid-low rectal cancer following curative surgery. *Psycho-Oncology*, 20, 706-714. doi:10.1002/pon.1763
- Smith, G. T., McCarthy, D. M., & Anderson, K. G. (2000). On the sins of short-form development. *Psychological Assessment*, 12, 102-111. doi:10.1037/1040-3590.12.1.102
- Stanton, J. M., Sinar, E. F., Balzer, W. K., & Smith, P. C. (2002). Issues and strategies for reducing the length of self-report scales. *Personnel Psychology*, 55, 167-194. doi:10.1111/j.1744-6570.2002.tb00108.x
- Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, 3, 4-70. doi:10.1177/109442810031002
- WHOQOL Group (1994). The development of the World Health Organization Quality of Life Assessment instrument (the WHOQOL). In J. Orley & W. Kuyken (Eds.), *Quality of life assessment: International perspectives* (pp. 41-57). Berlin, Germany: Springer-Verlag.

- WHOQOL Group (1995). Position paper from the World Health Organization. *Social Science and Medicine*, 41, 1403-1409. doi:10.1016/0277-9536(95)00112-K
- Wool, C., Cerutti, R., Marquis, P., Cialdella, P., & Hervie, C. (2000). Psychometric validation of two Italian quality of life questionnaires in menopausal women. *Maturitas*, 35, 129-142. doi:10.1016/S0378-5122(00)00093-1

APPENDIX A

The Psychological General Well-Being Index-A (PGWBI-A) (English Version)

Anxiety (Item Q8)

Were you generally tense or did you feel any tension during the past month?

- 0 Yes, extremely tense, most or all of the time
- 1 Yes, very tense most of the time
- 2 Not generally tense, but did feel fairly tense several times
- 3 I felt a little tense a few times
- 4 My general tension level was quite low
- 5 I never felt tense or any tension at all

Depression (Item Q3)

Did you feel depressed during the past month?

- 0 Yes, to the point that I felt like taking my life
- 1 Yes, to the point that I did not care about anything
- 2 Yes, very depressed almost every day
- 3 Yes, quite depressed several times
- 4 Yes, a little depressed now and then
- 5 No, never felt depressed at all

Positive well-being (Item Q20)

I felt cheerful, lighthearted during the past month.

- 0 None of the time
- 1 A little of the time
- 2 Some of the time
- 3 A good bit of the time
- 4 Most of the time
- 5 All of the time

Self-control (Item Q4r)

Have you been in firm control of your behavior, thoughts, emotions or feelings during the past month?

- 0 Yes, definitely so
- 1 Yes, for the most part
- 2 Generally so
- 3 Not too well
- 4 No, and I am somewhat disturbed
- 5 No, and I am very disturbed

General health (Item Q10r)

Did you feel healthy enough to carry out the things you like to do or had to do during the past month?

- 0 Yes, definitely so
- 1 For the most part
- 2 Health problems limited me in some important ways
- 3 I was only healthy enough to take care of myself
- 4 I needed some help in taking care of myself
- 5 I needed someone to help me with most or all of the things I had to do

Vitality (Item Q6r)

How much energy, pep, or vitality did you have or feel during the past month?

- 0 Very full of energy, lots of pep
- 1 Fairly energetic most of the time
- 2 My energy level varied quite a bit
- 3 Generally low in energy or pep
- 4 Very low in energy or pep most of the time
- 5 No energy or pep at all, I fell drained, sapped

The Psychological General Well-Being Index-B (PGWBI-B) (English Version)

Anxiety (Item Q19r)

Did you feel relaxed, at ease or high strung, tight, or keyed-up during the past month?

- 0 Felt relaxed and at ease the whole month
- 1 Felt relaxed and at ease most of the time
- 2 Generally felt relaxed but at times felt fairly high strung
- 3 Generally felt high strung but at times felt fairly relaxed
- 4 Felt high strung, tight, or keyed-up most of the time
- 5 Felt high strung, tight, or keyed-up the whole month

Depression (Item Q11)

Have you felt so sad, discouraged, hopeless, or had so many problems that you wondered if anything was worthwhile during the past month?

- 0 Extremely so, to the point that I have just about given up
- 1 Very much so
- 2 Quite a bit
- 3 Some, enough to bother me
- 4 A little bit
- 5 Not at all

Positive well-being (Item Q9r)

How happy, satisfied, or pleased have you been with your personal life during the past month?

- 0 Extremely happy, could not have been more satisfied or pleased
- 1 Very happy most of the time
- 2 Generally satisfied, pleased
- 3 Sometimes fairly happy, sometimes fairly unhappy
- 4 Generally dissatisfied or unhappy
- 5 Very dissatisfied or unhappy most or all the time

Self-control (Item Q18)

I was emotionally stable and sure of myself during the past month.

- 0 None of the time
- 1 A little of the time
- 2 Some of the time
- 3 A good bit of the time
- 4 Most of the time
- 5 All of the time

General health (Item Q13)

Have you been concerned, worried, or had any fears about your health during the past month?

- 0 Extremely so
- 1 Very much so
- 2 Quite a bit
- 3 Some, but not a lot
- 4 Practically never
- 5 Not at all

Vitality (Item Q21r)

I felt tired, worn out, used up, or exhausted during the past month.

- 0 None of the time
- 1 A little of the time
- 2 Some of the time
- 3 A good bit of the time
- 4 Most of the time
- 5 All of the time

Note. r = reverse item.

APPENDIX B

The Psychological General Well-Being Index-A (PGWBI-A) (Italian Version)

Anxiety (Item Q8)

Nelle ultime quattro settimane, è stato generalmente teso o ha provato tensione?

- 0 Sì, sono stato estremamente teso per tutto o quasi tutto il tempo
- 1 Sì, sono stato molto teso per la maggior parte del tempo
- 2 Generalmente no, ma mi è successo diverse volte di sentirmi piuttosto teso
- 3 Alcune volte mi sono sentito un po' teso
- 4 Il mio livello di tensione è stato piuttosto basso
- 5 Non ho mai avuto la sensazione di essere teso

Depression (Item Q3)

Nelle ultime quattro settimane, si è sentito depresso?

- 0 Sì, al punto di pensare di farla finita
- 1 Sì, al punto che non mi importava più di nulla
- 2 Sì, mi sono sentito molto depresso quasi tutti i giorni
- 3 Sì, mi sono sentito piuttosto depresso parecchie volte
- 4 Sì, mi sono sentito un po' depresso qualche volta
- 5 No, non mi sono mai sentito depresso

Positive well-being (Item Q20)

Nelle ultime quattro settimane, mi sono sentito allegro e sereno.

- 0 Mai
- 1 Quasi mai
- 2 Una parte del tempo
- 3 Molto tempo
- 4 Quasi sempre
- 5 Sempre

Self-control (Item Q4r)

Nelle ultime quattro settimane, si è sentito padrone delle Sue situazioni, pensieri, emozioni e dei Suoi sentimenti?

- 0 Sì, senz'altro
- 1 Sì, quasi del tutto
- 2 Sì, generalmente
- 3 Non troppo
- 4 No, e questo mi disturba un po'
- 5 No, e questo mi disturba molto

General health (Item Q10r)

Nelle ultime quattro settimane, si è sentito così bene da fare quello che desiderava o doveva fare?

- 0 Sì, decisamente
- 1 Sì, per fare quasi tutto quello che desideravo o dovevo fare
- 2 I miei problemi di salute mi hanno limitato in alcune cose importanti
- 3 A causa della mia salute sono appena in grado di prendermi cura di me stesso
- 4 Ho avuto bisogno di qualche aiuto per occuparmi di me stesso
- 5 Ho avuto bisogno di aiuto per tutto o quasi tutto quello che dovevo fare

Vitality (Item Q6r)

Nelle ultime quattro settimane, quanta energia, vivacità, o vitalità ha avuto o ha sentito di avere?

- 0 Decisamente pieno di energia, molto vivace
- 1 Abbastanza pieno di energia per la maggior parte del tempo
- 2 Ho avuto notevoli alti e bassi di vitalità ed energia
- 3 Il mio livello di energia o vitalità è stato generalmente basso
- 4 Il mio livello di energia o vitalità è stato quasi sempre molto basso
- 5 Mi sono sentito senza forze, svuotato, privo di energia o vitalità

The Psychological General Well-Being Index-B (PGWBI-B) (Italian Version)

Anxiety (Item Q19r)

Nelle ultime quattro settimane, si è sentito rilassato, tranquillo, oppure si è sentito molto teso, nervoso o agitato?

- 0 Sempre rilassato e tranquillo
- 1 Quasi sempre rilassato e tranquillo
- 2 Generalmente rilassato e tranquillo, ma qualche volta abbastanza teso
- 3 Generalmente molto teso, ma qualche volta abbastanza rilassato
- 4 Quasi sempre molto teso, nervoso, o agitato
- 5 Sempre molto teso, nervoso, o agitato

Depression (Item Q11)

Nelle ultime quattro settimane, si è sentito tanto triste, scoraggiato, disperato o ha avuto tanti problemi da chiedersi se valesse la pena andare avanti?

- 0 Sì, enormemente, tanto da essere quasi sul punto di lasciar perdere tutto
- 1 Sì, moltissimo
- 2 Sì, parecchio
- 3 Sì, abbastanza, tanto da turbarmi
- 4 Un po'
- 5 Per nulla

Positive well-being (Item Q9r)

Nelle ultime quattro settimane, in che misura si è sentito felice, soddisfatto o contento della Sua vita personale?

- 0 Veramente molto felice, non mi sarei potuto sentire più soddisfatto o contento
- 1 Quasi sempre molto felice
- 2 In generale molto soddisfatto, contento
- 3 A volte abbastanza felice, a volte piuttosto infelice
- 4 In generale insoddisfatto o infelice
- 5 Quasi sempre o sempre molto insoddisfatto o infelice

Self-control (Item Q18)

Nelle ultime quattro settimane, mi sono sentito emotivamente stabile e sicuro di me stesso.

- 0 Mai
- 1 Quasi mai
- 2 Una parte del tempo
- 3 Molto tempo
- 4 Quasi sempre
- 5 Sempre

General health (Item Q13)

Nelle ultime quattro settimane, ha provato apprensione, preoccupazione o paura per la Sua salute?

- 0 Enormemente
- 1 Moltissimo
- 2 Parecchio
- 3 Un po', ma non tanto
- 4 Quasi mai
- 5 Per nulla

Vitality (Item Q21r)

Nelle ultime quattro settimane, mi sono sentito stanco, esaurito, logorato o sfinito.

- 0 Mai
- 1 Quasi mai
- 2 Una parte del tempo
- 3 Molto tempo
- 4 Quasi sempre
- 5 Sempre

Note. r = reverse item.
