

## Psychometric Properties of the Brazilian 12-Item Short-Form Health Survey Version 2 (SF-12v2)

Bruno Figueiredo Damásio<sup>1</sup>  
Universidade Federal do Rio de Janeiro, Rio de Janeiro-RJ, Brazil

Thiago Francisco Andrade  
Universidade Federal da Paraíba,  
João Pessoa-PB, Brazil

Sílvia Helena Koller  
Universidade Federal do Rio Grande do Sul, Porto Alegre-RS, Brazil

**Abstract:** The 12-Item Short-Form Health Survey, in its initial (SF-12) and revised form (SF-12v2) is a widely used measure to evaluate health-related quality of life (HRQoL). The present study evaluates the factor structure and reliability of the Brazilian version of the SF-12v2. Participants were 627 subjects (74.1% women), aged from 18 to 88 years ( $M = 38.6$ ;  $SD = 13.16$ ), from 17 Brazilian states. Confirmatory factor analyses suggested two pairs of error terms to be highly correlated (3a-3b; and 4a-4b). A qualitative inspection showed an overlap of content among these items. The respecified model presented adequate fit indices. Convergent validity was also tested with measures of health-related self-care, subjective happiness, life satisfaction, depression and self-efficacy. Expected correlations were found between the SF-12v2 and these measures. Results showed initial evidence in favor of using the SF-12v2 as a measure of physical and mental health in the Brazilian context.

**Keywords:** health, factor analysis, psychometrics

### Propriedades Psicométricas da 12-Item Short-Form Health Survey Versão 2 (SF-12v2)

**Resumo:** O instrumento 12-Item Short-Form Health Survey, em sua versão inicial (SF-12) e revisada (SF-12v2), é uma escala amplamente utilizada na avaliação da qualidade de vida relacionada à saúde (QVRS). O presente estudo investiga a estrutura fatorial e confiabilidade da versão brasileira do SF-12v2. Participaram do estudo 627 sujeitos (74,1% mulheres), com idades entre 18 e 88 anos ( $M = 38,6$ ;  $DP = 13,16$ ), oriundos de 17 Estados brasileiros. Análises fatoriais confirmatórias apresentaram uma alta correlação entre dois pares de erros (itens 3a-3b e 4a-4b). Uma inspeção qualitativa sugeriu sobreposição de conteúdo entre esses itens. O modelo reespecificado apresentou adequados índices de ajuste. A validade convergente foi investigada com medidas de autocuidados relacionados à saúde, felicidade subjetiva, satisfação com a vida, depressão e autoeficácia. Foram encontradas correlações esperadas entre a SF-12v2 e essas medidas. Os resultados mostram evidências iniciais favoráveis ao uso do SF-12v2 como medida de saúde física e saúde mental no contexto brasileiro.

**Palavras-chave:** saúde, análise fatorial, psicometria

### Propiedades Psicométricas de la 12-Item Short-Form Health Survey Versión 2 (SF-12v2)

**Resumen:** El 12-Item Short-Form Health Survey, en su versión inicial (SF-12) y revisada (SF-12v2) es un instrumento ampliamente utilizado en la evaluación de la calidad de vida relacionada a la salud (CVRS). Este estudio evalúa la estructura factorial y fiabilidad de la versión brasileña del SF-12v2. Participaron 627 sujetos (74,1% mujeres), con edad entre 18 y 88 años ( $M = 38.6$ ;  $DE = 13.16$ ), provenientes de 17 estados brasileños. Análisis factoriales confirmatorios presentaron elevadas correlaciones entre dos pares de errores (ítems 3a-3b y 4a-4b). Una inspección cualitativa sugirió una superposición de contenido entre ellos. El modelo reestructurado presentó índices de ajuste adecuados. La validez convergente fue investigada con medidas de autocuidados relacionadas a la salud, felicidad subjetiva, satisfacción con la vida, depresión e autoeficacia. Fueron encontradas correlaciones esperadas entre la SF-12v2 y esas medidas. Los resultados muestran evidencias iniciales favorables para la SF-12v2 como una medida de salud física y mental en el contexto brasileño.

**Palabras clave:** salud, análisis factorial, psicometría

The 36-Item Short-Form Health Survey (SF-36) and its shorter version, the SF-12, are the measures most widely used to evaluate health-related quality of life (HRQoL). These instruments have been translated into more than 140 languages (Burholt & Nash, 2011). Among other reasons, their widespread use is linked to the fact that the

instruments are general health survey measures, which can be used in both clinical (Yarlas et al., 2011) and non-clinical samples (Burholt & Nash, 2011), and in a wide variety of age groups, from adults 18 years of age (Hopman et al., 2009) to elderly people (Lima et al., 2009).

The SF-36 factorial structure encompasses eight scales: 1) physical functioning (PF); 2) role limitations due to physical health (RP); 3) bodily pain (BP); 4) general health (GH); 5) vitality (VT); 6) social functioning (SF); 7) role limitations due to emotional problems (RE); and 8) mental health (MH). These eight scales are condensed into two larger dimensions, the Physical Component Summary (PCS),

<sup>1</sup> Correspondence address:

Bruno Figueiredo Damásio. Instituto de Psicologia / Departamento de Psicometria – Universidade Federal do Rio de Janeiro (UFRJ). Avenida Pasteur, 250. Urca. CEP 22290-902. Rio de Janeiro-RJ, Brazil. E-mail: bf.damasio@gmail.com

including the PF, RP, BP, and GH factors and the Mental Component Summary (MCS), including the VT, SF, RE, and MH factors (Ware, Snow, Kosinski, & Gandek, 1993).

In order to develop a shorter instrument, which could reliably reproduce the PCS and the MCS of the SF-36, Ware, Kosinski and Keller (1996) designed the SF-12. The 12 items of the SF-12 were derived from data from a US general population survey in 1990 (Ware et al., 1996). In that study, the authors used regression methods to select the most significant items of the SF-36, which could reliably reproduce the PCS and the MCS explained variance of the SF-36. Thus, the SF-12 comprises items from all the eight scales of the original SF-36, as it was designed to predict the PCS and MCS as close as possible to the original 36-item structure. The resulting 12-item short-form (SF-12) achieved multiple R-squared values of .91 and .92 in predictions of the SF-36 PCS and SF-36 MCS scores, respectively (Ware et al., 1996). These results have been replicated in several other European countries. In 1998, for example, Gandek et al. (1998), in a cross-validation study, tested the standard 12-item selection suggested in the original U.S study (Ware et al., 1993) for nine European countries (Denmark, France, Germany, Italy, the Netherlands, Norway, Spain, Sweden, and the United Kingdom). The authors found that the SF-12 items explained a high level of the variance of the PCS-36 and MCS-36 scores (89-92% and 88-94%, respectively) across all nine countries. These results provided considerable evidence of concurrent validity for the SF-12 in relation to the SF-36.

A second version of the SF-12, entitled SF-12v2 (Ware, Kosinski, Turner-Bowker, & Gandek, 2002), is also available. The SF-12v2 incorporated several modifications in relation to the standard SF-12. Improvements in the instructions, items and layout of the questionnaire were performed aiming to simplify the wording and make the instrument less ambiguous. Because of the common ceiling and/or floor effects, some items had their response options exchanged from a dichotomous (yes/no) to a 5-point scale and other items had their response options reduced from a 6-point to a 5-point scale. Despite the improvements, a large part of the instrument was unchanged, and its essentials (use, factor structure and analysis) remained the same (Ware et al., 2002).

Regarding the factor structure of the SF-12 (in the standard and second version), some studies have found that the proposed bi-factorial structure (PCS and MCS) presented acceptable fit indices (Kontodimopoulos, Pappa, Niakas, & Tountas, 2007; Montazeri, Vahdaninia, Mousavi, & Omidvari, 2009; Okonkwo, Roth, Pulley, & Howard, 2010). However, some studies found that the SF-12 presents factorial problems. In some studies, acceptable fit indices for the SF-12 could not be easily achieved. For example, Maurischat, Ehlebracht-König, Kühn, and Bullinger (2006), in testing the factorial validity of the SF-12 in a German

sample survey of patients with inflammatory-rheumatic disease ( $N = 545$ ), could find acceptable fit indices only by allowing several items to cross-load in more than one factor or by allowing error correlations. Similar results were found in the Jakobsson, Westergren, Lindskov, and Hagell (2012) study. These authors evaluated the factor structure of the SF-12 in three Swedish samples (elderly people aged +75,  $N = 4,278$ ; people with Parkinson's disease,  $N = 159$ ; and stroke survivors,  $N = 89$ ). Exploratory factor analyses, using Parallel Analysis as the factor retention method, failed to support the bi-factorial structure among elderly people and stroke survivors, suggesting a three-factor model as the most reliable to the data. When the bi-factorial solution was forced, cross-loadings of the SF-12 items in all three samples were found. Confirmatory factor analyses also presented lack of fit between empirical data and the original proposed model. In part, these problems are remnants from the problems evidenced in the SF-36 factor structure (for a review, see Güthlin & Walach, 2007; Keller et al., 1998; Vet, Adèr, Terwee, & Pouwer, 2005), as the development study of the SF-12 did not consider any factorial method in the item retention procedures.

Regarding Brazil, only the standard version of the SF-12 is available, and it has been used in several empirical studies (Abreu, Walker, Sesso, & Ferraz, 2011; Aquino et al., 2009; Martins, Polvero, Rocha, Foss, & Santos Junior, 2012). However, all studies that have employed the Brazilian standard SF-12 as a HRQoL measure used the previously established factor structure of the instrument, without evaluating its plausibility neither by exploratory nor confirmatory factor analyses.

Considering that the factor structure is one of the most important issues of construct validity in psychological research, and considering the fact that the SF-12 (in both its original and revised versions, SF-12v2) is one of the instruments most used to evaluate HRQoL worldwide, the present study aimed to present the Brazilian version of the SF-12v2 and to evaluate its factor structure.

## Method

### Participants

Participants were 627 subjects (74.1% women), aged from 18 to 88 years ( $M = 38.6$ ;  $SD = 13.16$ ), from 17 Brazilian states. A total of 37.6% were married, 25.4% single, 14.5% cohabitating, 12.6% dating or engaged, 7.8% divorced, 0.8% widowed, and 1.3% in other situations (not specified). The sample was composed of participants who took part in a larger study entitled "Subjective Well-Being, Orientations to Happiness and its Implications in Psychological Well-Being and Health-Related Self-Care: Adaptations of Questionnaires and Psychosocial Investigations", which aims to evaluate personal and contextual factors related to positive psychological functioning.

## Instruments

*Sociodemographic Questionnaire.* This gathered information about gender, age, religious practice (presence or absence), educational level, financial income, job satisfaction, and other sociodemographic information.

*12-item Short-Form Health Survey – Version 2 (SF-12v2).* This is a self-report measure of health-related quality of life (HRQoL) designed to investigate multidimensional aspects of physical and mental health for the general population and those with chronic diseases. Two summary subscales may be derived from the SF-12, including a mental health summary and a physical health summary. The Brazilian version employed in this study was based on the Brazilian adaptation of the SF-36 (Ciconelli, Ferraz, Santos, Meinão, & Quaresma, 1999). The 12 items that compose the SF-12 (Ware et al., 2002) were chosen, and the required modifications of the first version to the revised version (SF-12v2) were conducted. Specifically, the following changes were implemented: Response scale from items 3a, 3b, 4a and 4b were changed from dichotomous (yes or no) to ordered polytomous categories (All of the time; Most of the time; Some of the time; A little of the time; None of the time); the response scale from items 6a, 6b and 6c were changed from a 6-point scale (All of the time; Most of the time; A good bit of the time; Some of the time; A little of the time; None of the time) to a 5-point scale (All the time; Most of the time; Some of the time; A little of the time; Never). These suggestions were in accordance with the original English version of the SF-12v2 (Ware et al., 2002). The final version of the scale is shown in Appendix A.

*Subjective Happiness Scale (SHS).* The SHS is a 4-item test that evaluates happiness from the respondent's own perspective. The instrument has presented excellent psychometric properties in several countries (Moghnie & Kazarian, 2012; Spagnoli, Caetano, & Silva, 2012; Shimai, Otake, Utsuki, Ikemi, & Lyubomirsky, 2004; Swami, 2008; Swami et al., 2009). In the validation study (Lyubomirsky & Lepper, 1999), the authors found adequate reliability index, with alpha coefficients varying from .80 to .94 in 14 different samples ( $N = 2,732$ ). The Brazilian version of the SHS was validated by Damásio, Zanon, and Koller (2014), and presented excellent fit indices: CFI = 1.00; TLI = 1.02; SRMR = .006; RMSEA (90% CI) = .000 (.000 - .006). In this study, the goodness-of-fit indices of the SHS were: CFI = 1.00; TLI = 1.00; RMSEA (90% CI) = .01 (.000 - .072); SRMR = .01.

*Satisfaction with Life Scale (SWLS).* The Brazilian version of the SWLS was adapted and validated by Gouveia, Milfont, Fonseca, and Coelho (2009). The instrument is composed of five items, which evaluate life-satisfaction from a subjective perspective (e.g., "In general, I am satisfied with my life"). In the validation study, the scale presented adequate psychometric properties (Reliability index,  $\alpha = .80$ ; goodness-of-fit indices: GFI = .99; TLI = .98; CFI = .99;

RMSEA = .06; SRMR = .02). In this study, the goodness-of-fit indices of the SWLS were: CFI = 1.00; TLI = 1.00; RMSEA = .01 (.00 - .06); SRMR = .01.

*12-item General Health Questionnaire (GHQ-12).* The GHQ-12 (Goldberg et al., 1997) is the reduced version of the original General Health Questionnaire (Goldberg, 1972) and is one of the instruments most widely used to evaluate indicators of psychological well-being. Responses are given on a scale ranging from 1 (more than habitual) to 4 (less than habitual). Brazilian validation studies reported a two-factor solution (depression and self-efficacy) as the most reliable, with Alpha reliability index ranging from .85 to .63 (Damásio, Machado, & Silva, 2011; Gouveia, Barbosa, Andrade, & Carneiro, 2010; Sarriera, Schwarcz, & Câmara, 1996). In this study, the goodness-of-fit of this bifactorial solution (self-efficacy and depression) were: CFI = .98; TLI = .98; RMSEA (90% CI) = .07 (.060 - .082); SRMR = .05.

*Appraisal of Self-Care Agency Scale-Revised (ASAS-R).* The ASAS-R is a 15-item measure that evaluates the level of self-care agency using a five-point Likert scale ranging from 1 (totally disagree) to 5 (totally agree). It is composed of three factors (having capacity for self-care; developing capacity for self-care; lacking capacity for self-care). The Brazilian version of the instrument (Damásio & Koller, 2013) presented adequate fit indices and reliability. In this study, goodness-of-fit indices of the expected three factor solution were: CFI: .97; TLI: .96; RMSEA (90% CI) = .06 (.05 - .07).

## Procedure

**Data collection.** Participants were accessed through different sources: Personal and media invitations, and snowball technique (Patton, 1990). Those who decided to participate answered a web-based survey.

**Data analysis.** Using Mplus 6.11, we performed a confirmatory factor analysis with the co-variance matrix as input, using the robust MLM estimation method (Satorra & Bentler, 2001). The hypothesized 2-factor structure (Physical and Mental Health, Model 1) was tested and the goodness-of-fit was assessed by means of the following fit indices: Satorra-Bentler scaled  $\chi^2$ , CFI, TLI, and RMSEA. Modification indices were evaluated in order to seek model misspecification. According to guidelines (Brown, 2006), the  $\chi^2$  value must be non-significant, providing evidence that the observed matrix is not significantly different to the population matrix. The RMSEA value must be less than .06 or .08 (with higher-bound 90% confidence interval not exceeding .10). The CFI and TLI values should be greater than .90 (preferably greater than .95) (Brown, 2006).

In order to evaluate the convergent validity for the SF-12v2, we correlated the scores of the physical and mental health factors with the scores of several constructs which theoretically comprises the nomothetic network of the SF-12v2, namely: subjective happiness, life satisfaction, general health, and health-related self-care (having, developing and lacking the capacity for self-care).

**Ethical Considerations**

This study was approved by the Institutional Review Board (Ethics Committee) of the Universidade Federal do Rio Grande do Sul (Protocol no. 22240/2012). The terms of consent form was placed on the first page of the survey so that participants could only advance in the questionnaire by accepting the terms and consenting to participate in the study.

**Results**

**Confirmatory Factor Analysis**

Initially, we sought to examine the theoretical bi-dimensional structure of the SF-12v2 by means of a confirmatory factor analysis (CFA). As can be seen in Table 1, the overall fit of Model 1 to the data was poor. Chi-square was significant, although this measure is strongly affected by sample size. However, the remaining fit indices were out of the

commonly adopted parameters of good fit (CFI = .80; TLI = .78; RMSEA = .11), which also suggested inadequacy of the model.

The inspection of modification indices showed that imposing a co-variance between the error terms associated to the pairs of items 3a-3b and 4a-4b, respectively, would improve model fit. These items refer to the limitations in daily activities as a result of physical or mental health problems. As the content of these items overlap each other, we tested the fit of a respecified model, imposing the two aforementioned co-variances among the error terms associated to those items. Table 1 shows the fit indices of the respecified model (Model 2). As in the first model,  $\chi^2$  remained significant. However, the other fit indices show that this respecified model presented an acceptable fit, CFI = .937, TLI = .918, RMSEA = .066 (90% CI = .056 - .076).

Table 2 shows the standardized estimates for Model 2. All the estimates were significantly different from zero, and the saturation values ranged from .681 to -.448 in the Physical Health factor, and from .764 to .652 in the Mental Health

Table 1  
*Goodness-of-Fit Indices for the Initial and Respecified Models*

Model	$\chi^2$	CFI	TLI	RMSEA (90% CI)	BIC
Initial	(53) 434.4326*	.799	.779	.107 (.098-.117)	17193.734
Respecified	(51) 187.483*	.937	.918	.066 (.056-.076)	16883.871

Note. CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; RMSEA = Root Mean Square Error of Approximation.  
\* $p < .001$ .

Table 2  
*Factor Structure With Standardized Estimates for the Respecified Model*

Factor	Item	Estimate	SE	CR*
Physical Health	1. General health rating (Avaliação geral da saúde)	.520	.036	14.571
	2a. Moderated activities (Atividades moderadas)	-.492	.051	-9.734
	2b. Climb several flights (Subir vários lances de escada)	-.448	.049	-9.180
	3a. Accomplished less tasks than would like (Realizou menos tarefas do que gostaria)	-.535	.041	-13.018
	3b. Limited in kind of work or other activities (Esteve limitado no seu tipo de trabalho ou outras atividades)	-.563	.041	-13.796
	5. Pain interferes in daily tasks (Dor interferiu nas atividades diárias)	.681	.038	17.698
Mental Health	4a. Accomplished less than would like (Realizou menos tarefas do que gostaria)	.655	.030	21.593
	4b. Did work or activities less carefully than usual (Não realizou atividades com o cuidado habitual)	.652	.029	22.129
	6a. Felt Peaceful (Tem se sentido calmo ou tranquilo?)	-.702	.028	-24.799
	6b. Felt Energetic (Tem se sentido com muita energia?)	-.708	.032	-21.786
	6c. Felt Blue/Sad (Tem se sentido desanimado e abatido?)	.764	.025	30.603
	7a. Problems interfered in social activities (Problemas interferiram nas atividades sociais)	.698	.028	25.054

Note. Estimate refers to factor loadings in the confirmatory model; SE = Standard Error; CR = Critical Ratio. In parenthesis, abbreviated content of the items presented in Brazilian-Portuguese language.  
\* $p < .001$ .

factor. The estimated correlation between the two factors was .645. Alpha reliability was adequate for both subscales (.73 for the PH, and .86 for the MH).

**Correlations With Other Relevant Measures**

Table 3 presents the convergent validity for the SF-12v2. As can be seen, the physical health factor showed significant negative correlations with lack of capacity for self-care ( $r = -.35, p < .001$ ) and the depression GHQ subscale ( $r = -.25, p < .001$ ), and positive correlations with satisfaction with life ( $r = .31, p < .001$ ), the self-efficacy GHQ subscale ( $r = .30, p < .001$ ), subjective happiness ( $r = .28, p < .001$ ), having capacity for self-care ( $r = .25, p < .001$ ), and developing capacity for self-care ( $r = .14, p < .001$ ).

A similar pattern was found for the mental health factor, with significant correlations with the self-efficacy GHQ subscale ( $r = .53, p < .001$ ), the depression GHQ subscale ( $r = -.52, p < .001$ ), subjective happiness and life satisfaction ( $r = .48, p < .001$ ), lack of capacity for self-care ( $r = -.39, p < .001$ ), having capacity for self-care ( $r = .31, p < .001$ ) and developing capacity for self-care ( $r = .20, p < .001$ ).

**Discussion**

In this study, we aimed to evaluate the factorial structure of the Brazilian version of the SF-12v2, and to present evidence of convergent validity by assessing the relationships between the SF-12v2 subscales and several measures that compose its nomothetic network. To date, no evidence has been presented for the psychometric robustness of the Brazilian SF-12v2, regarding both factorial-structural and construct-related aspects.

The results of the confirmatory factor analysis showed that the items of the SF-12v2 loaded as expected in the PHC

and MHC subscales, ranging from .52 for item 1 to .76 for item 6c. Alpha reliabilities were also adequate and similar to previous studies (Cheak-Zamora, Wyrwich, & McBride, 2009; De Smedt et al., 2013; Montazeri et al., 2011). Although the factor loadings and alpha reliability were adequate, a confirmatory factor analysis evaluation of the SF-12v2 provided non-acceptable fit indices. These results were consonant with those found by Maurischat et al. (2006), which showed that the SF-12v2 had acceptable fit indices only when cross-loadings and error term correlations were added. In the present study, the modification indices did not suggest the presence of cross-loadings, although the imposition of co-variances between two pairs of error terms were carried out in order to improve fit. Our respecified model showed acceptable fit indices, in a similar way to Montazeri et al. (2009). A qualitative inspection of the items in which error terms were correlated (3a with 3b; and 4a with 4b) showed high overlapping content, referring to limitations in daily activities as a result of physical (items 3a and 3b) and mental health (items 4a and 4b) issues (see Table 2). Item overlap content is a known source of model misfit, leading to the need for imposing additional parameters (in this case, co-variances between error terms) on the model, though these model modifications need to be soundly theoretically justified (Byrne, 2009; Tabachnick & Fidell, 2013), as in this case. According to these findings, it is possible to argue in favor of a reduced and more parsimonious version of the SF-12v2. In this study, we opted to maintain the scale with its original structure because of the nature of our non-representative and non-clinical sample. It is necessary to evaluate whether these problems continue in future studies. If so, arguments toward the refinement of the SF-12v2 become more convincing.

Regarding the convergent validity, we found evidence that corroborated the SF-12v2 suitability. The mental and physical health subscales were positively and moderately associated. This result corroborates previous findings suggesting the interconnection between physical and mental health, and more precisely, between the PCS and MHS factors of the SF scales (Hobart, Williams, Moran, & Thompson, 2002; Montazeri et al., 2011; Simon, Revicki, Grothaus, & Vonkorff, 1998). Both PCS and MCS subscales also presented adequate convergent validity with the external measures (SHS, SWLS, GHQ-12, and ASAS-R). Correlations of the expected magnitude and direction were found. It was possible to identify that the MCS presented stronger correlations with the employed measures, while the PCS presented weaker correlations. This result was expected, as our convergent measures were more closely associated with mental health than with physical health.

This article presents advances and limitations that must be highlighted. Initially, we present the psychometric properties and the Brazilian version of the SF-12v2, a globally employed measure to evaluate health-related quality of life (HRQoL). This version incorporates all the modifications previously conducted in the original American

Table 3  
*Pearson's Correlations Between SF-12v2 Subscales and Other Relevant Measures (N = 627)\**

	SF-12 Physical Health	SF-12 Mental Health
SHS	.28	.48
SWLS	.31	.48
GHQ - Depression**	-.25	-.52
GHQ - Self-efficacy**	.30	.53
ASAS - Having	.25	.31
ASAS - Developing	.14	.20
ASAS - Lacking	-.35	-.39

Note. \*\*N = 613; PH = Physical Health; MH = Mental Health; GHQ 1 and GHQ 2 = General Health Questionnaire Depression and Self-efficacy subscales; SHS = Subjective Happiness Scale; SWLS = Satisfaction with Life Scale; ASAS-Having, ASAS-Developing, and ASAS-Lacking represent Having, Developing, and Lacking subscales of The Appraisal of Self-Care Agency Scale-Revised (ASAS-R).

\* $p < .001$ .

version (Ware et al., 2002), and updates the version of the scale to the Brazilian context. In terms of limitations, the participants (mostly women) composed a non-representative sample. Additionally, the present study did not include a clinical sample, therefore the results cannot be generalized to other contexts. So far, the results indicate the plausibility of using the SF-12v2 as a reliable measure to evaluate physical and mental health in non-clinical populations. Further studies are required to extend the sources of validity of the Brazilian version of the SF-12v2.

## References

- Abreu, M. M., Walker, D. R., Sesso, R. C., & Ferraz, M. B. (2011). Health-related quality of life of patients receiving hemodialysis and peritoneal dialysis in São Paulo, Brazil: A longitudinal study. *Value in Health, 14*(5, Suppl. 1), S119-121. doi:10.1016/j.jval.2011.05.016
- Aquino, N. M. R., Sun, S. Y., Oliveira, E. M., Martins, M. G., Silva, J. F., & Mattar, R. (2009). Sexual violence and its association with health self-perception among pregnant women. *Revista de Saúde Pública, 43*(6), 954-960. doi:10.1590/S0034-89102009005000068
- Burholt, V., & Nash, P. (2011). Short Form 36 (SF-36) Health Survey Questionnaire: Normative data for Wales. *Journal of Public Health, 33*(4), 587-603. doi:10.1093/pubmed/fdr006
- Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York, NY: Guilford.
- Byrne, B. M. (2009). *Structural equation modeling with AMOS: Basic concepts, applications, and programming* (2nd ed.). New York, NY: Routledge.
- Cheak-Zamora, N. C., Wyrwich, K. W., & McBride, T. D. (2009). Reliability and validity of the SF-12v2 in the medical expenditure panel survey. *Quality of Life Research, 18*(6), 727-735. doi:10.1007/s11136-009-9483-1
- Ciconelli, R. M., Ferraz, M. B., Santos, W., Meinão, I., & Quaresma, M. R. (1999). Tradução para a língua portuguesa e validação do questionário genérico de avaliação de qualidade de vida SF-36 (Brasil SF-36). *Revista Brasileira de Reumatologia, 39*(3), 143-150.
- Damásio, B. F., & Koller, S. H. (2013). The Appraisal of Self-Care Agency Scale - Revised (ASAS-R): Adaptation and construct validity in the Brazilian context. *Cadernos de Saúde Pública, 29*(10), 2071-2082. doi:10.1590/0102-311X00165312
- Damásio, B. F., Machado, W. L., & Silva, J. P. (2011). Estrutura fatorial do Questionário de Saúde Geral (QSG-12) em uma amostra de professores escolares. *Avaliação Psicológica, 10*(1), 99-105.
- Damásio, B. F., Zanon, C., & Koller, S. H. (2014). Validation and psychometric properties of the Brazilian version of the Subjective Happiness Scale. *Universitas Psychologica, 13*(1), 17-24. doi:10.11144/Javeriana.UPSY13-1.vppb
- De Smedt, D., Clays, E., Doyle, F., Kotseva, K., Prugger, C., Pająk, A., ... De Bacquer, D. (2013). Validity and reliability of three commonly used quality of life measures in a large European population of coronary heart disease patients. *International Journal of Cardiology, 167*(5), 2294-2299. doi:10.1016/j.ijcard.2012.06.025
- Gandek, B., Ware, J. E., Jr., Aaronson, N. K., Alonso, J., Apolone, G., Bjorner, J., ... Sullivan, M. (1998). Tests of data quality, scaling assumptions, and reliability of the SF-36 in eleven countries: Results from the IQOLA Project. International Quality of Life Assessment. *Journal of Clinical Epidemiology, 51*(11), 1149-1158. doi:10.1016/S0895-4356(98)00106-1
- Goldberg, D. P. (1972). *The detection of psychiatric illness by questionnaire*. London, United Kingdom: Oxford University Press.
- Goldberg, D. P., Gater, R., Sartorius, N., Ustun, T. B., Piccinelli, M., Gureje, O., & Rutter, C. (1997). The validity of two versions of the GHQ in the WHO study of mental illness in general health care. *Psychological Medicine, 27*(1), 191-197.
- Gouveia, V. V., Barbosa, G. A., Andrade, E. O., & Carneiro, M. B. (2010). Factorial validity and reliability of the General Health Questionnaire (GHQ-12) in the Brazilian physician population. *Cadernos de Saúde Pública, 26*(7), 1439-1445. doi:10.1590/S0102-311X2010000700023
- Gouveia, V. V., Milfont, T. L., Fonseca, P. N., & Coelho, J. A. P. M. (2009). Life satisfaction in Brazil: Testing the psychometric properties of the Satisfaction With Life Scale (SWLS) in five Brazilian samples. *Social Indicators Research, 90*(2), 267-277. doi:10.1007/s11205-008-9257-0
- Güthlin, C., & Walach, H. (2007). MOS-SF 36: Structural equation modelling to test the construct validity of the second-order factor structure. *European Journal of Psychological Assessment, 23*(1), 15-23. doi:10.1027/1015-5759.23.1.15
- Hobart, J. C., Williams, L. S., Moran, K., & Thompson, A. J. (2002). Quality of life measurement after stroke: Uses and abuses of the SF-36. *Stroke, 33*(5), 1348-1356. doi:10.1161/01.STR.0000015030.59594.B3
- Hopman, W. M., Berger, C., Joseph, L., Towheed, T., Prior, J. C., Anastassiades, T., ... CaMos Research Group. (2009). Health-related quality of life in Canadian adolescents and young adults: Normative data using the SF-36. *Canadian Journal of Public Health, 100*(6), 449-452.
- Jakobsson, U., Westergren, A., Lindskov, S., & Hagell, P. (2012). Construct validity of the SF-12 in three different samples. *Journal of Evaluation in Clinical Practice, 18*(3), 560-566. doi:10.1111/j.1365-2753.2010.01623.x

- Keller, S. D., Ware, J. E., Jr., Bentler, P. M., Aaronson, N. K., Alonso, J., Apolone, G., ... Gandek, B. (1998). Use of structural equation modeling to test the construct validity of the sf-36 health survey in ten countries: Results from the IQOLA Project. *Journal of Clinical Epidemiology*, *51*(11), 1179-1188. doi:10.1016/S0895-4356(98)00110-3
- Kontodimopoulos, N., Pappa, E., Niakas, D., & Tountas, Y. (2007). Validity of SF-12 summary scores in a Greek general population. *Health Quality of Life Outcomes*, *5*, 55. doi:10.1186/1477-7525-5-55
- Lima, M. G., Barros, M. B. A., César, C. L. G., Goldbaum, M., Carandina, L., & Ciconelli, R. M. (2009). Health related quality of life among the elderly: A population-based study using SF-36 survey. *Cadernos de Saúde Pública*, *25*(10), 2159-2167. doi:10.1590/S0102-311X2009001000007
- Lyubomirsky, S., & Lepper, H. S. (1999). A measure of subjective happiness: Preliminary reliability and construct validation. *Social Indicators Research*, *46*(2), 137-155. doi:10.1023/A:1006824100041
- Martins, M. R. I., Polvero, L. O., Rocha, C. E., Foss, M. H., & Santos Junior, R. D. (2012). Using questionnaires to assess the quality of life and multidimensionality of fibromyalgia patients. *Revista Brasileira de Reumatologia*, *52*(1), 21-26. doi:10.1590/S0482-50042012000100003
- Maurischat, C., Ehlebracht-König, I., Kühn, A., & Bullinger, M. (2006). Factorial validity and norm data comparison of the Short Form 12 in patients with inflammatory-rheumatic disease. *Rheumatology International*, *26*(7), 614-621. doi:10.1007/s00296-005-0046-7
- Moghnie, L., & Kazarian, S. S. (2012). Subjective happiness of Lebanese college youth in Lebanon: Factorial structure and invariance of the Arabic subjective happiness scale. *Social Indicators Research*, *109*(2), 203-210. doi:10.1007/s11205-011-9895-5
- Montazeri, A., Vahdaninia, M., Mousavi, S. J., Asadi-Lari, M., Omidvari, S., & Tavousi, M. (2011). The 12-item medical outcomes study short form health survey version 2.0 (SF-12v2): A population-based validation study from Tehran, Iran. *Health and Quality of Life Outcomes*, *9*, 12. doi:10.1186/1477-7525-9-12
- Montazeri, A., Vahdaninia, M., Mousavi, S. J., & Omidvari, S. (2009). The Iranian version of 12-item Short Form Health Survey (SF-12): Factor structure, internal consistency and construct validity. *BMC Public Health*, *9*, 341. doi:10.1186/1471-2458-9-341
- Okonkwo, O. C., Roth, D. L., Pulley, L., & Howard, G. (2010). Confirmatory factor analysis of the validity of the SF-12 for persons with and without a history of stroke. *Quality of Life Research*, *19*(9), 1323-1331. doi:10.1007/s11136-010-9691-8
- Patton, M. Q. (1990). *Qualitative evaluation and research methods* (2nd ed.). Newbury Park, CA: Sage.
- Sarriera, J. C., Schwarcz, C., & Câmara, S. G. (1996). Bem-estar psicológico: Análise fatorial da escala de Goldberg (GHQ-12) numa amostra de jovens. *Psicologia: Reflexão e Crítica*, *9*(2), 293-306.
- Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*, *66*(4), 507-514. doi:10.1007/BF02296192
- Shimai, S., Otake, K., Utsuki, N., Ikemi, A., & Lyubomirsky, S. (2004). Development of a Japanese version of the Subjective Happiness Scale (SHS), and examination of its validity and reliability. *Nihon Koshu Eisei Zasshi*, *51*(10), 845-853. doi:10.11236/jph.51.10\_845
- Simon, G. E., Revicki, D. A., Grothaus, L., & Vonkorff, M. (1998). SF-36 summary scores: Are physical and mental health truly distinct? *Medical Care*, *36*(4), 567-572.
- Spagnoli, P., Caetano, A., & Silva, A. (2012). Psychometric properties of a Portuguese version of the Subjective Happiness Scale. *Social Indicators Research*, *105*(1), 137-143. doi:10.1007/s11205-010-9769-2
- Swami, V. (2008). Translation and validation of the Malay Subjective Happiness Scale. *Social Indicators Research*, *88*(2), 347-353. doi:10.1007/s11205-007-9195-2
- Swami, V., Stieger, S., Voracek, M., Dressler, S. G., Eisma, L., & Furnham, A. (2009). Psychometric evaluation of the Tagalog and German Subjective Happiness Scales and a cross-cultural comparison. *Social Indicators Research*, *93*(2), 393-406. doi:10.1007/s11205-008-9331-7
- Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Boston, MA: Pearson.
- Vet, H. C. W., Adèr, H. J., Terwee, C. B., & Pouwer, F. (2005). Are factor analytical techniques used appropriately in the validation of health status questionnaires? A systematic review on the quality of factor analysis of the SF-36. *Quality of Life Research*, *14*(5), 1203-1218. doi:10.1007/s11136-004-5742-3
- Ware, J. E., Jr., Snow, K. K., Kosinski, M., & Gandek, B. (1993). *SF-36 Health Survey manual and interpretation guide*. Boston, MA: New England Medical Center/The Health Institute.
- Ware, J. E., Jr., Kosinski, M., & Keller S. D. (1996). A 12-item Short-Form Health Survey: Construction of scales and preliminary tests of reliability and validity. *Medical Care*, *34*(3), 220-233.
- Ware, J. E., Jr., Kosinski, M., Turner-Bowker, D. M., & Gandek, B. (2002). *How to score version 2 of the SF-12 Health Survey*. Lincoln, RI: QualityMetric.
- Yarlas, A. S., White, M. K., Yang, M., Saris-Baglama, R. N., Bech, P. G., & Christensen, T. (2011). Measuring the health status burden in hemodialysis patients using the SF-36® health survey. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care & Rehabilitation*, *20*(3), 383-389. doi:10.1007/s11136-010-9764-9768

*Bruno Figueiredo Damásio* is an Assistant Professor of the Instituto de Psicologia at the Universidade Federal do Rio de Janeiro.

*Thiago Francisco Andrade* is a Ph.D. candidate of the Graduate Program in Social Psychology at the Universidade Federal da Paraíba.

*Sílvia Helena Koller* is a Full Professor of the Instituto de Psicologia at the Universidade Federal do Rio Grande do Sul.

*Received:* June 3, 2014

*1st Revision:* Sep. 1, 2014

*2nd Revision:* Oct. 29, 2014

*Approved:* Dec. 2, 2014

*How to cite this article:*

Damásio, B. F., Andrade, T. F., & Koller, S. H. (2015). Psychometric properties of the Brazilian 12-Item Short-Form Health Survey Version 2 (SF-12v2). *Paidéia (Ribeirão Preto)*, 25(60), 29-37. doi: 10.1590/1982-43272560201505

APPENDIX A

**Medical Outcomes 12-Item Short-Form Health Survey (Versão 2) – SF-12v2**

Este questionário busca compreender a sua opinião em relação à sua saúde. Essas informações irão ajudar a avaliar como você se sente e o quão bem você está em relação às suas atividades diárias. Por favor, responda cada pergunta selecionando a resposta mais apropriada. Se você não tiver certeza sobre como responder à pergunta, por favor, dê a resposta que mais se aproxima do que você pensa.

1. Em geral, você diria que sua saúde é:

1	2	3	4	5
Excelente	Muito Boa	Boa	Ruim	Muito Ruim

2. Os seguintes itens são sobre atividades que você poderia fazer atualmente durante um dia comum. **Devido à sua saúde**, você tem dificuldade para fazer essas atividades? Neste caso, quanto?

ATIVIDADES	Sim. Dificulta muito.	Sim. Dificulta um pouco.	Não. Não dificulta de modo algum.
a. <b>Atividades moderadas</b> , tais como mover uma mesa, passar aspirador de pó, jogar bola, varrer a casa.	1	2	3
b. Subir <b>vários</b> lances de escada.	1	2	3

3. Durante as **últimas 4 semanas**, quanto do tempo você teve algum dos seguintes problemas com seu trabalho ou com alguma atividade diária regular, **como consequência de sua saúde física**?

	Todo o tempo	A maior parte do tempo	Alguma parte do tempo	Uma pequena parte do tempo	Nenhuma parte do tempo
a. Realizou menos tarefas do que você gostaria?	1	2	3	4	5
b. Esteve limitado no seu tipo de trabalho ou outras atividades?	1	2	3	4	5

4. Durante as **últimas 4 semanas**, quanto do tempo você teve algum dos seguintes problemas com seu trabalho ou outra atividade regular diária, **como consequência de algum problema emocional** (por exemplo, sentir-se deprimido ou ansioso)?

	Todo o tempo	A maior parte do tempo	Alguma parte do tempo	Uma pequena parte do tempo	Nenhuma parte do tempo
Realizou menos tarefas do que você gostaria?	1	2	3	4	5
Não trabalhou ou não fez qualquer das atividades com tanto cuidado como geralmente faz?	1	2	3	4	5

5. Durante as **últimas 4 semanas**, quanto a dor interferiu com seu trabalho normal (incluindo tanto o trabalho, fora de casa e dentro de casa)?

De maneira alguma	Um pouco	Moderadamente	Bastante	Extremamente
1	2	3	4	5

6. Estas questões são sobre como você se sente e como tudo tem acontecido com você durante as **últimas 4 semanas**. Para cada questão, por favor, dê uma resposta que mais se aproxime da maneira como você se sente. Em relação **às últimas 4 semanas**:

	Todo o tempo	A maior parte do tempo	Alguma parte do tempo	Uma pequena parte do tempo	Nenhuma parte do tempo
a. Quanto tempo você tem se sentido calmo ou tranquilo?	1	2	3	4	5
b. Quanto tempo você tem se sentido com muita energia?	1	2	3	4	5
c. Quanto tempo você tem se sentido desanimado e abatido?	1	2	3	4	5

7. Durante as **últimas 4 semanas**, quanto do seu tempo a sua saúde física ou problemas emocionais interferiram com suas atividades sociais (como visitar amigos, parentes, etc.)?

Todo o tempo	A maior parte do tempo	Alguma parte do tempo	Uma pequena parte do tempo	Nenhuma parte do tempo
1	2	3	4	5