

## **VALIDATION OF THE BALANCED MEASURE OF PSYCHOLOGICAL NEEDS (BMPN) IN SPANISH AND PORTUGUESE: METHOD EFFECTS ASSOCIATED TO NEGATIVELY WORDED ITEMS**

Laura Galiana<sup>1</sup>, Melchor Gutiérrez<sup>1</sup>, José M. Tomás<sup>1</sup>,  
and Patricia Sancho<sup>2</sup>

<sup>1</sup>*University of Valencia;* <sup>2</sup>*Catholic University San Antonio de Murcia (Spain)*

### **Abstract**

The self-determination theory is one of the most popular theories on motivation. It postulates that three basic needs mediate humans' proactivity and involvement: autonomy, competence, and relatedness. Among the measurement instruments addressing these needs, the Balanced Measure of Psychological Needs (BMPN) has been recently presented. This research aims to study the dimensionality of the BMPN, in its Portuguese and Spanish versions. Samples of 2034 and 715 students were recruited from high school classes in Angola and the Dominican Republic, respectively. The better-fitting model for the Portuguese and Spanish versions of the BMPN scale was the one posing a method artifact. This evidence, together with lower factor loadings, problems with reliability, and poorer nomological validity for the negatively worded items, moved us to propose a positive version of the BMPN, with better psychometric properties. Our results support latest research criticism on the harmful effects of including negatively worded items in scale development.

KEY WORDS: *basic psychological needs, self-determination theory, measurement artifacts, motivation, structural equation modeling.*

### **Resumen**

La teoría de la autodeterminación es una de las más populares en motivación. Postula que tres necesidades básicas median la proactividad y el compromiso de las personas: autonomía, competencia y relación. De entre los instrumentos de medida que evalúan estas necesidades, la "Medida balanceada de necesidades psicológicas" (*Balanced Measure of Psychological Needs*, BMPN) se ha presentado recientemente. Esta investigación tiene como objetivo estudiar la dimensionalidad de la BMPN, en sus versiones portuguesa y española. Se recogieron muestras de 2034 y 715 estudiantes en clases de instituto en Angola y República Dominicana, respectivamente. El modelo que mejor ajustó en las versiones portuguesa y española de la escala BMPN fue el que incluyó un factor de método. Esta evidencia, junto con las bajas saturaciones factoriales, problemas con la fiabilidad y una pobre validez nomológica para los ítems invertidos, nos

llevó a proponer una versión positiva de la BMPN, con mejores propiedades psicométricas. Nuestros resultados apoyan las críticas que investigaciones recientes hacen sobre los efectos dañinos de incluir ítems invertidos en el desarrollo de escalas.

*PALABRAS CLAVE: necesidades psicológicas básicas, teoría de la autodeterminación, artefactos de medida, motivación, modelos de ecuaciones estructurales.*

## Introduction

Theories on motivation have greatly expanded during the last decades. Among them, the self-determination theory (SDT) (Deci & Ryan, 1985, 2000) has been one of the most blossoms. During more than 30 years of history, the SDT has served as a starting point for science advance in several psychological areas, i.e. work (Deci et al., 2001), sport (Adie, Duda, & Ntoumanis, 2012), relationships (Sheldon & Bettencourt, 2002), aggression (de Haan, Soenens, Dekovic, & Prinzie, 2013), or school psychology (Filak & Sheldon, 2003; Niemiec, Ryan, & Deci, 2009). Among the five motivational 'minitheories' (Reeve, 2012) included in the SDT, this paper is centered on the basic needs theory (BNT). The BNT assumes our basic psychological needs are three (autonomy, competence and relatedness), and they need to be satisfied (Deci & Ryan, 1985, 2000). Autonomy is referred to self-regulation; competence is defined as the experience of mastery; and relatedness is conceptualized as the experience of support and connection (Deci & Ryan, 2000). The fulfillment of these three basic needs will entail wellbeing (Deci & Ryan, 2000; Reis, Sheldon, Gable, Roscoe, & Ryan, 2000; Ryan, 1995; Sheldon & Niemiec, 2006; Vansteenkiste & Ryan, 2013).

Measurement of needs satisfaction has been done with several methods, in various settings, and within several approaches (Johnston & Finney, 2010; Sheldon & Hilpert, 2012). Several scales have been designed for specific contexts (Brien et al., 2012; Vlachopoulos & Michailidou, 2006; Wilson, Rogers, Rodgers, & Wild, 2006), some others are domain-specific (i. e., Weinstein, Przybylski, & Ryan, 2012), and others are more general, centered on the three needs defined by the SDT. Within all them, the most used and representative one is the Basic Psychological Needs Scale (BPNS; Gagné, 2003), with almost 200 citations in the Web of Knowledge in the last 10 years. The BPNS assesses the basic psychological needs with 21 items, and has been used both to measure needs satisfaction in general (with a total score) and the three basic needs (with three different scores or subscales) (Johnston & Finney, 2010).

Whereas a vast amount of research has focused on this needs fulfillment, their measurement has been traditionally undertaken under the assumption of adequate psychometric properties never proved, and only lately researchers have paid attention to its adequacy. When construct validity of the BPNS was first tested, Johnston and Finney (2010) hypothesized two structural models for the totality of the scale, a one-factor and a three-factor solution. Results of these models showed no appropriate fit and, thus, some changes on the scale were carried on. The authors tested several models with response-bias factor associated

to the negatively worded items. A 16-item three-factor model with method effect was championed in both samples (Johnston & Finney, 2010). Together with these difficulties on the dimensionality of the scale, problems such as low reliabilities associated with the autonomy and the competence dimensions, the large amount of non-explained variance (Johnston & Finney, 2010), the unbalance in the number of items per scale, or the presence of items which might assess multiple needs at the same time (Sheldon & Hilpert, 2012) have been pointed out.

In order to address this quandary, Sheldon and Hilpert (2012) presented the Balanced Measure of Psychological Needs (BMPN). This scale assesses the basic needs referred by Deci and Ryan (1985, 2000) by means of 18 statements: six indicators per psychological need, three positively worded and three negatively worded, and it was previously used in Sheldon and Gunz (2009) and Sheldon, Abad, and Hinsch (2011). In the validation work, however, some controversy arose with its dimensionality. Neither a single-factor structure nor a three-factor solution fitted appropriately the data. Meanwhile, a model with five factors, the three general need factors and two factors explaining needs satisfaction and dissatisfaction, produced the best fit (Sheldon & Hilpert, 2012).

Although Sheldon and Hilpert's (2012) approach was based on the substantive interpretation of the method factors found in both scales (named satisfaction and dissatisfaction of needs), the truth is that there is a long tradition in the self-esteem research arena explaining these as method artifacts produced by the negative wording of some items (Bachman & O'Malley, 1986; Bagozzi, 1993; Carmines & Zeller, 1974, 1979; Corwyn, 2000; Dalal & Carter, 2015; Goldsmith, 1986; Hensley & Roberts, 1976; Kaufman, Rasinski, Lee, & West, 1991; Kohn, 1977; Marsh, 1996; Marsh & Grayson, 1995; Salgado & Iglesias, 1995; Tomás & Oliver, 1999, 2004; Tomás, Oliver, Galiana, Sancho, & Lila, 2013; Wang, Siegal, Falck, & Carlson, 2001). In this line of thought, some studies have shed light on these phenomena, by explaining method effects as a response style (De Jonge & Slaets, 2005; DiStefano & Motl, 2006; Quilty, Oakman, & Risko, 2006). It seems of primary importance, then, to elucidate if the dimensions found in these scales correspond or not to method artifacts, because, as it is widely known, these artifacts can affect psychometric characteristics of the scales and distort, among others, association measures, failing to show the true relationships among the variables of interest (Tomás, Meléndez, Oliver, Navarro, & Zaragoza, 2010).

However, the confusion in the factor structure of the BMPN could be explained by other complexities in the data rather than response bias or method effects. Recently, Reise, Morizot and Hays (2007) and Reise (2012) have "rediscovered" the utility of a subtype of structural model, the bifactor model to understand the complex structure of some psychological measures. Bifactor models specify that the covariance among a set of item responses can be accounted for by a single general factor that reflects the common variance running among all scale items, and domain-specific factors reflecting additional common variance among items. Such a structure would consist of a model with a latent variable underlying all the items (basic psychological needs), and three uncorrelated domain-specific basic needs: relatedness, competence, and autonomy.

As it is clear from the aforementioned, further research on the basic psychological needs assessment is needed. Specifically, the aim of this research is to study the dimensionality of the BMPN. This research is based on the validation of the BMPN presented in American university students by Sheldon and Hilpert (2012), while adding some novelties. Firstly, to our knowledge, there is no other adaptation or validation of the BMPN neither in English nor in other languages. Current research offers a validation of both the Spanish and Portuguese versions of the scale. Secondly, main factor structures tested in Sheldon and Hilpert (2012) were analyzed, but adding two new models. A model with the three substantive factors (relatedness, competence and autonomy), plus a method factor only associated to the negatively worded items has been specified. This model was based on previous literature that points out that negatively worded items produce larger bias than those positively worded. Additionally, a completely substantive bifactor model has also been added to those studied by the authors. This bifactor model decomposes the items' variance into a general trait component (basic psychological needs) and domain-specific traits, namely relatedness, competence and autonomy. Thirdly, this research offers a detailed study on the consequences of using both positive and negatively worded items on criterion-related validity for this scale. Fourthly, a new version of the BMPN is proposed, together with its psychometric properties.

## Method

### *Samples*

*Angolan sample.* A total of 2034 students studying seventh to twelfth grades in Benguela province (Angola) were sampled. Their mean age was 17.5 years old ( $SD= 2.31$ ). 50.1% were women ( $n= 1035$ ). 52.8% lived in urban areas. They were sampled in their school settings. The sampling scheme was non-probabilistic. The researchers tried to recruit schools of different social and economic backgrounds (and types: private, public, religious...) to enter the samples, but no probabilistic selection was made. The survey was self-administered, but interviewers were present and solved any doubts or hesitations the participants had. Almost all participants completed the survey, but there were a few (less than 1%) students who did not consistently answer all parts of the survey and their questionnaires were not considered.

*Dominican Republic sample.* A total of 715 students studying middle education in Santo Domingo, the capital of the Dominican Republic were sampled. The sampling scheme was probabilistic, with schools chosen based on a stratification by type of school (public, private and polytechnics) in order to represent the population of students of middle education in Santo Domingo (error rate 5%, with confidence of 95%). Their mean age was 15.5 years old ( $SD= 1.57$ ). 58% were women. With respect to the type of school, most students went to public schools (58.3%), enrolled in private schools there were 22.8% of the students, and the remaining 18.8% of the students were attending polytechnics

(schools that offer professional education). Again the response rate was more than 99%. The procedure was the same as the one employed in Angola.

### *Instruments*

The *Balanced Measure of Psychological Needs* (BMPN; Sheldon & Hilpert, 2012) was used. This scale assesses the three basic needs of the SDT (Deci & Ryan, 1985, 2000): relatedness, competence, and autonomy. Example items are, respectively: "I felt a sense of contact with people who care for me and who I care for"; "I was successfully completing difficult tasks and projects"; and "I was free to do things my own way". It is composed of 18 items, six for each dimension, three negatively worded. A deeper description can be consulted in the Introduction section. The rating scale was from 1 (*totally disagree*) to 5 (*totally agree*).

Additionally, two questionnaires of engagement the Student Engagement in School Scale (SES; Veiga, 2013) and the Engagement Scale by Nie and Lau (2009) were used to assess criterial (nomological) validity of the BMPN. The SES evaluates four dimensions of students' engagement with school: cognitive, affective, behavioral, and agentic engagement, through 20 items, five per dimension. The original scale presented in Portuguese (Veiga, 2013) was used in the study in Angola, with internal consistencies of: .62, .66, .76, and .62 for the cognitive, affective, behavioral, and agentic engagement. An adaptation into Spanish was used in the study in Dominican Republic, with reliability coefficients of .67, .74, .84, and .65 for the four dimensions of engagement. Nie and Lau (2009) scale of engagement evaluates a single dimension of engagement by means of five items with internal consistencies of .73 in the Angolan sample and .71 in the Dominican Republic sample. Both scales had a rating scale from 1 (*totally disagree*) to 5 (*totally agree*).

### *Statistical analyses*

A set of competing confirmatory factor models (CFA) were specified, estimated and tested in Mplus 7.3 (Muthén & Muthén, 2007). According to the ordinal nature of the data and its non-normality weighted least square mean and variance (WLSMV corrected estimation) estimation was used. Several criteria were used to assess goodness-of-fit: (a) the chi-square statistic; (b) the comparative fit index (CFI); and (c) the root mean squared error of approximation (RMSEA). A model with a CFI of .95 or larger and a RMSEA of .08 or lower would be indicative of very good fit between the hypothesized model and the data (Hu & Bentler, 1999).

The models to compare were nested. When nested models are compared there are two rationales (Little, 1997), the statistical and the modeling one. The statistical approach employs  $\chi^2$  differences ( $\Delta\chi^2$ ) to compare constrained to unconstrained models, with non-significant values supporting the more parsimonious model. This statistical approach has been criticized (Cheung & Rensvold, 2002; Little, 1997), recommending the modeling approach that uses

practical fit indices to determine the overall adequacy of a fitted model. From this point of view, if a parsimonious model evinces adequate levels of practical fit, then it is preferred over the more complex model. Usually, CFI differences ( $\Delta\text{CFI}$ ) are used to evaluate measurement invariance. CFI differences lower than .01 (Cheung & Rensvold, 2002) or .05 (Little, 1997) are usually employed as cut-off criteria.

Reliability of the dimensions in the study was estimated using Cronbach's alphas and composite reliability indexes (CRI). Cronbach's coefficient alpha is the most widely used estimator of internal consistency of tests. Nevertheless, it has been criticized as being only completely appropriate with essentially tau-equivalent items (and tests), and also by being a lower bound for the true reliability (Raykov, 2004). More explicitly, a tau-equivalent test assumes all items measure the same latent variable, on the same scale, with the same degree of precision, with all true scores being equal (Graham, 2006). When tau-equivalence does not hold, alpha will over or underestimate (more often the latter) the population value. An alternative to coefficient alpha is the CRI, which is usually calculated using estimates from confirmatory factor analyses (Graham, 2006). Accordingly, both alpha, as a popular measure of internal consistency, and the more adequate CRI, were calculated.

In sum, the statistical procedures to test for psychometric properties followed the usual guidelines used, for example, in Fernández-Castillo, Vílchez-Lara, and Sada-Lázaro (2012), Faria and Lima-Santos (2012), or Caballo *et al.* (2010).

## Results

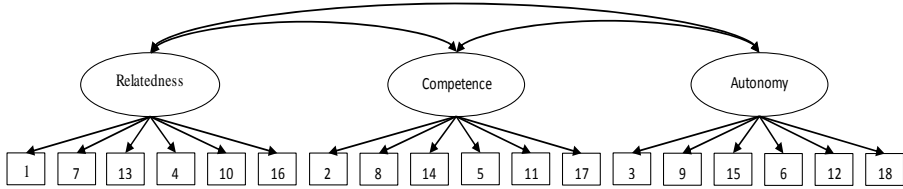
### *Factorial validity*

Four competitive CFAs were specified and estimated. These models may be seen in Figure 1. The four models were tested in both samples:

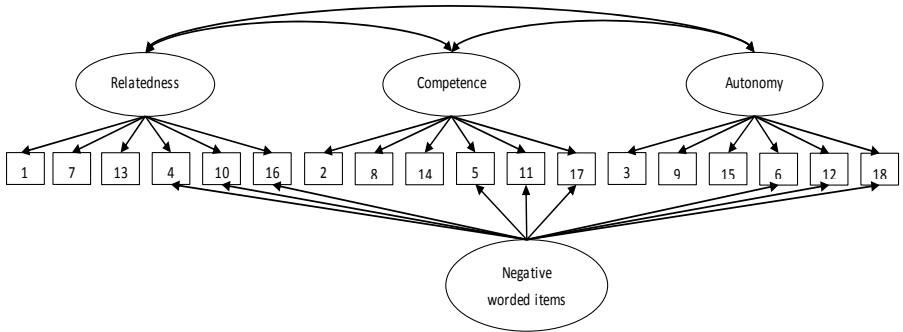
- Model 1 is the three-factor substantive structure that guided the development of the scale, with three correlated factors: relatedness, competence, and autonomy.
- Model 2 adds a method factor associated to negatively worded items to the three-factor structure. This response bias is assumed to be uncorrelated to the substantive latent variables. As seen in the introduction, this model was based on the bulk of existing literature that points at the importance of method effects associated to negatively worded items.
- Model 3 posited the same three substantive factors, plus two correlated method factors associated, respectively, to negative and positively worded items. This model was the best-fitting model in the original version of the scale (Sheldon & Hilpert, 2012).
- Model 4 is a bifactor model. This model decomposes items' variance into a general trait component (basic psychological needs) and domain-specific traits, namely relatedness, competence and autonomy. It has never been tested on this scale.

**Figure 1**  
Hypothesized models

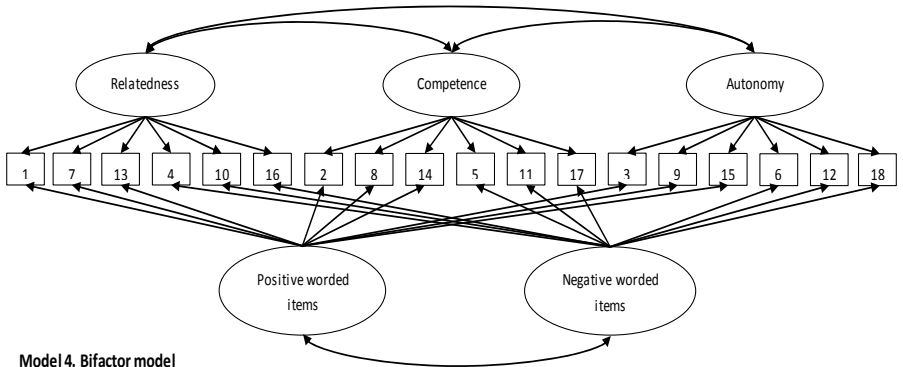
**Model 1. Three-factor model**



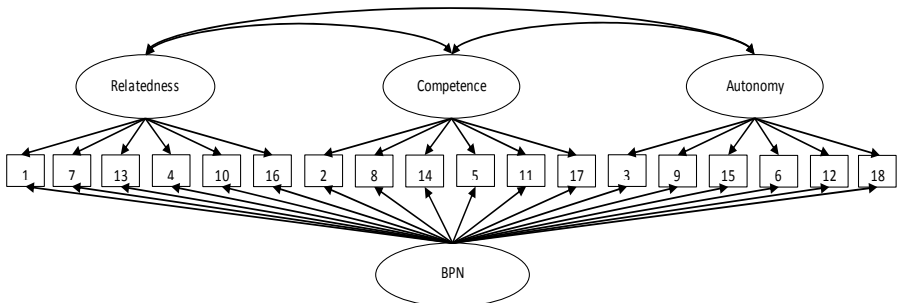
**Model 2. Correlated traits plus a method factor**



**Model 3. Correlated traits correlated methods**



**Model 4. Bifactor model**



Goodness-of-fit indices are shown in Table 1 for both samples. Models 2 and 3 fitted the data well, whereas model 1 and 4 were ill-fitted. That held true for both samples. Overall, this result means that adding wording bias into the substantive factors is needed in order to achieve a good fit. However, the fit indices of models 2 and 3 were very close, and it was therefore needed to compare them. Model 2 was more parsimonious as it only introduced a method factor effect due to negatively worded items. When chi-square differences tests were used, there were statistically significant differences favoring model 3 both in Angola ( $\Delta\chi^2= 66.038$ ,  $\Delta df= 10$ ,  $p < .001$ ) and Dominican Republic ( $\Delta\chi^2= 77.26$ ,  $\Delta df= 10$ ,  $p < .001$ ). However, from a practical fit perspective, the differences between both models in both samples were negligible: the  $\Delta CFI$  for the Angolan sample was .009, and the  $\Delta CFI$  for Dominican Republic sample was .011. Additionally, a careful look at the standardized solutions for both models made clear that method effects associated to negatively worded items were much larger than those associated to positively worded. Therefore, model 2 was retained. Standardized factor loadings for the best-fitting models can be consulted in Table 2.

**Table 1**  
Goodness of fit indices for the tested models

Samples/Models	$\chi^2$	<i>df</i>	<i>p</i>	CFI	RMSEA
Angola sample					
Three-factor model	2327.696	132	< .001	.674	.091
Correlated traits plus a method factor	519.649	123	< .001	.941	.040
Correlated traits correlated methods	452.346	113	< .001	.950	.038
Bifactor model	1457.397	117	< .001	.801	.075
Dominican Republic sample					
Three-factor model	3009.683	132	< .001	.472	.175
Correlated traits plus a method factor	419.602	123	< .001	.946	.058
Correlated traits correlated methods	318.723	113	< .001	.962	.050
Bifactor model	2635.696	117	< .001	.538	.174



**Table 2**  
Standardized factor loadings for the best-fitting models

Item	Correlated traits plus a method factor						Correlated traits correlated methods											
	Relatedness		Competence		Autonomy		Negative		Relatedness		Competence		Autonomy		Negative		Positive	
	ANG	DOR	ANG	DOR	ANG	DOR	ANG	DOR	ANG	DOR	ANG	DOR	ANG	DOR	ANG	DOR	ANG	DOR
1	.59	.76							.57	.67							.17	.33
2			.42	.64							.36	.54					.21	.35
3					.35	.24							.12	.28			.51	.02 ns
4	-.10	-.26					.46	.54	-.20	-.14						.45	.58	
5			.08	-.09			.45	.70			.00 ns	.05 ns				.45	.70	
6					.13	-.04 ns	.44	.68					.06 ns	.10	.45	.67		
7	.63	.70							.57	.66							.23	.22
8			.53	.61							.50	.56					.21	.24
9					.48	.62							.40	.59			.21	.19
10	.08	.00 ns					.45	.53	.00 ns	.11					.45	.52		
11			.15	.06 ns			.48	.58			.07 ns	.19			.49	.55		
12					.36	.17	.32	.42					.32	.27	.37	.37		
13	.54	.51							.55	.66							.11	-.16 ns
14			.51	.67							.46	.58					.23	.34
15					.46	.65							.33	.54			.32	.37
16	.21	.08					.39	.46	.14	.19					.42	.44		
17			.45	.43			.13	.17			.47	.49			.19	.08 ns		
18					.19	-.06 ns	.43	.57					.12	.06 ns	.45	.57		

Note: ANG= Angola; DOR= Dominican Republic; Negatively worded items and their factor loadings are written in italics; ns= not statistically significant

### *Internal consistency*

Cronbach's alphas were calculated for the three dimensions of the BMPN, in both Angolan and Dominican Republic samples. Alphas were .36, .40, and .42, for the relatedness, competence, and autonomy factors, respectively, in the Portuguese version of the BMPN; and .46, .49, and .45 in the Spanish version. As previously stated, CRI was also calculated, as this index has been determined to be a better estimate of the true reliability (Raykov, 2004) and, additionally, it takes into account variance resulting from different sources (i.e., the method factor). Values of CRI for the Portuguese version were .43, .45, and .39, for the relatedness, competence, and autonomy dimensions, respectively. In the case of the Spanish version of the BMPN, CRI were .47, .52, and .32, for relatedness, competence, and autonomy. All in all, information on reliability pointed out poor consistency for the three subscales in the two samples.

### *Nomological validity*

In order to test for the nomological validity of the BMPN scale, school engagement, and its dimensions, were employed as criterion. In the educational psychology arena, the satisfaction of psychological needs has very frequently been related to school engagement. Among these contributions, the most important is the longitudinal evidence presented in Jang, Kim and Reeve (2012) that relates in a panel design teacher's support of autonomy, psychological needs and school engagement. Other recent reference that relates school engagement and the satisfaction of psychological needs is the one by Raufelder et al. (2014). These two recent references also introduce a bulk of other studies that support the close relationship among these constructs.

First, the correlations among the three basic needs and the engagement dimension in Nie and Lau (2009) engagement scale were estimated. These correlations were separately calculated for each basic psychological need measured with all the items (positive and negative), only positive and only negative. The correlations of relatedness with engagement in Angola for the total, positive and negative versions were, respectively: .169 ( $p < .01$ ), .259 ( $p < .01$ ), and  $-.004$  ( $p > .05$ ). The same correlations in the Dominican Republic were: .78 ( $p < .01$ ), .362 ( $p < .01$ ), and  $-.067$  ( $p > .05$ ). When competence is considered, the pattern of correlations in Angola was: .186 ( $p < .01$ ), .372 ( $p < .01$ ), and  $-.153$  ( $p < .05$ ). In the Dominican Republic: .372 ( $p < .01$ ), .573 ( $p < .01$ ), and .057 ( $p > .05$ ). Finally, the correlations of engagement with autonomy (total, positive and negative) in Angola were: .119 ( $p < .01$ ), .244 ( $p < .01$ ), and  $-.076$  ( $p < .01$ ). In the Dominican Republic, these correlations were: .209 ( $p < .01$ ), .375 ( $p < .01$ ), and  $-.026$  ( $p > .05$ ). Clearly, the nomological validity of the scale suffered when negatively worded items are used to measure the three basic psychological needs for both samples.

Second, and given the method factor found in the BMPN structure, the nomological validity was calculated separating the dimensions measured with positive and negatively worded items and then correlated with the four dimensions of engagement in the SES scale. In short, the relationships between psychological

needs and the criteria (engagement) were calculated for relatedness, autonomy and competence measured positively and, separately, negatively. All these correlations are presented in Table 3. The nomological validity seemed adequate, but only for the positively worded dimensions. Relationships among the BMPN dimensions and the criteria were larger when these dimensions were measured by positively worded items, and this was true no matter the criteria, with the exception of behavioral engagement, a criterion (only) measured by negatively worded items.

**Table 3**  
Nomological validity correlation coefficients

Criteria	Positive items			Negative items		
	Relatedness	Competence	Autonomy	Relatedness	Competence	Autonomy
Angola						
Cognitive engagement	.25	.31	.24	.00	.07	.07
Affective engagement	.28	.26	.23	.07	.07	.10
Behavioral engagement	.08	.09	.02 ns	.24	.10	.18
Agentic engagement	.27	.29	.24	.07	.10	.11
Dominican Republic						
Cognitive engagement	.34	.42	.31	.05 ns	.07	.09
Affective engagement	.33	.39	.37	.13	.02 ns	.08
Behavioral engagement	.10	.21	.12	.30	.21	.29
Agentic engagement	.34	.34	.29	.03ns	.14	.09

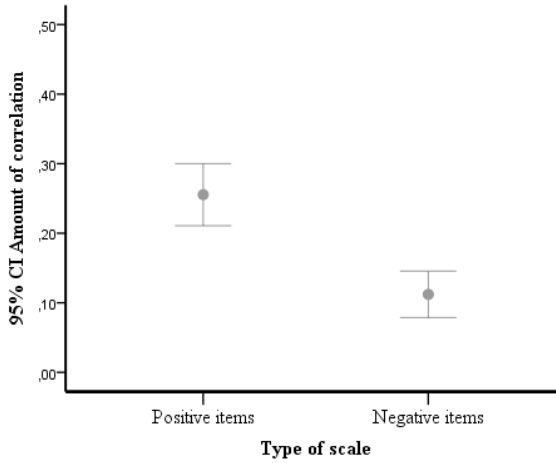
Note: all correlations statistically significant unless ns (non-significant) stated.

The correlations may be statistically analyzed in order to get a better insight about the impact of wording on the validity of the dimensions of psychological needs. Accordingly, correlations among the dimensions of engagement and the three dimensions of psychological needs, both measured positively and negatively, were compared by means of a *t*-test. Average correlation for the positively worded dimensions with the engagement factors was .25, whereas the average correlation with the negatively worded dimensions was .11,  $t(16) = 5.317$ ,  $p < .001$ ,  $d = 1.56$ . The mean correlations for both types of dimensions with their 95% confidence intervals are shown in Figure 2. When the distinction is made among the correlations with the four criteria, the results of the interaction in an ANOVA 2 (positively vs. negatively worded psychological needs) x 4 (engagement dimensions) also shown a statistically significant and large effect,

$F(3, 40) = 31.56, p < .001, \text{partial-}\eta^2 = 0.71$ . Means for the interaction are shown in Figure 3. It became apparent that correlations among the criteria and the positively worded psychological needs were larger with only one exception, behavioral engagement (the only engagement dimension with negatively worded items, indeed with all items negatively worded; misbehavior).

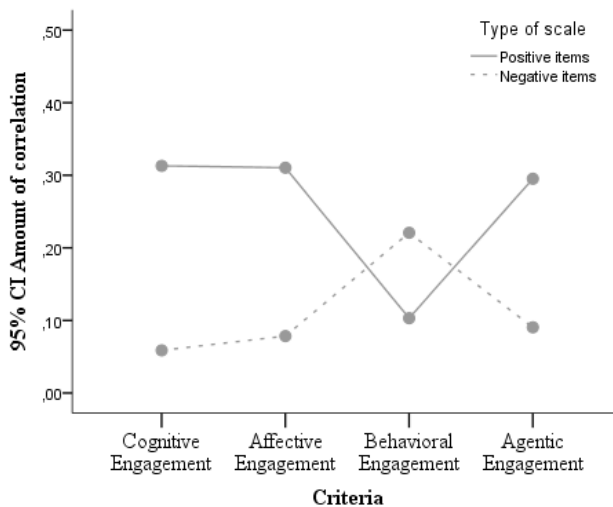
**Figure 2**

95% Confidence intervals for the mean correlation among the dimensions of the BMPN (positive and negative) and engagement



**Figure 3**

Means of amount of correlation by type of scale and criteria



*Positive version of the BMPN*

Taking into account the results of the Portuguese and Spanish versions of the BMPN hitherto reported, it seems pretty clear that the psychometric properties of the scale were not completely adequate, and therefore it needed to be reformulated for its appropriate usage. Thus, a new version of the scale was proposed: the Positive version of the Balanced Measure of Psychological Needs (see Appendix A), which retains only the three positively worded items per dimension.

In order to compare the relative psychometric properties of this positive version of the BMPN, all analyses were separately made for the positive version and the corresponding negative version of the BMPN. Firstly, two new CFAs were estimated and tested. Fit indices for both samples and both versions are shown in Table 4, and they were clearly better for the positive version of the BMPN and in each sample. Indeed the negative version had a very poor fit.

In the same line, reliability estimates were also better for the positive adaptation, with alphas of .53, .49, and .37 for the relatedness, competence, and autonomy dimensions of the Portuguese version, and of .61, .69, and .48, for the Spanish version, in front of alphas of .36, .27, and .39 of the negative Portuguese version, and .50, .43, and .48 for the Spanish version. Similar results were found when CRIs were calculated for both versions. In the positive versions, CRIs were .62, .55, and .44 for the Portuguese version, and .71, .69, and .53, for the Spanish one, whereas in the negative version, values of CRI were .41, .32, and .45 for the Portuguese version, and .51, .48, and .54 for the Spanish version.

**Table 4**

Goodness-of-fit indices for the positive and negative versions of the BMPN

BMPN	Angola					Dominican Republic				
	$\chi^2$	<i>df</i>	<i>p</i>	CFI	RMSEA	$\chi^2$	<i>df</i>	<i>p</i>	CFI	RMSEA
Positive version	66.713	24	< .001	.988	.030	66.603	24	< .001	.987	.050
Negative version	240.654	24	< .001	.913	.067	201.385	24	< .001	.926	.102

Note: BMPN= Balanced Measure of Psychological Needs.

**Discussion**

This research has presented psychometric results of a relatively new scale that measures the basic psychological needs, prominent dimensions in the self-determination theory: the Balanced Measure of Psychological Needs (BMPN; Sheldon & Hilpert, 2012). This new scale was originally presented in an English version and came to solve some psychometric problems associated to the most widely used scale to measure psychological needs, the BPNS (Gagné, 2003). These "problems" in the BPNS included lack of fit for the structural model, response bias associated to negatively worded items, low reliability associated to autonomy and competence dimensions, the unbalance in the number of indicators per factor, and

some indicators ambiguity (Johnston & Finney, 2010; Sheldon & Hilpert, 2012). However, when the BMPN scale, a balanced scale in terms of both items per factor and positive vs. negatively worded items, was validated (Sheldon & Hilpert, 2012), a five factor structure (which could be interpreted as method bias associated to item wording) aroused, and therefore the same “problems” that affected the BNPS were also present in the BMPN. Current research has extended the study on the psychometric properties of the BMPN in several ways: it has offered evidence on the psychometric properties of the BMPN in two languages different from the original English version, Spanish and Portuguese; it has analyzed reliability and validity in two relatively large samples of countries usually under-investigated; and it has tested more *a priori* structures than the initial validation of the scale.

In a nutshell, the analyses of both samples always agreed that the BMPN presented some methodological problems in terms of structure, validity and reliability, and that this was, in our opinion, largely due to the presence of response bias associated to negatively worded items. When Sheldon and Hilpert (2012) proposed a balanced measure, meaning that the same number of items should measure each dimension and that the same number of positively and negatively worded items should be employed within each dimension, they were following the advice of respected psychometricians such as Likert (1932), Cronbach (1942), Nunnally (1978), or Anastasi (1982), to name just a few. This has been for long the psychometric advice to avoid the negative effects of response-styles (acquiescence among others). Unfortunately, they were also following psychometric advice that can be proved to be a methodological myth and urban legend. Indeed, Dalal and Carter (2015) have clearly shown that the practice of introducing negatively worded items was based on four assumptions: a) including negatively worded items will avoid response bias (response styles or response sets); b) negative items will not impact scales’ quality; c) negatively worded items will not affect validity; and d) negatively and positively worded items measure the same construct. Unfortunately, Dalal and Carter (2015) also demonstrated that none of these assumptions holds true, and they do so based on a huge amount of analytical and empirical evidence. Our results also point in exactly the same direction.

With respect to the assumption that negatively worded items do not impact quality of the scales, this was not the case with the BMPN since all reliability estimates in both versions of the scale were markedly lower when negatively worded items were included. This held true for both alpha and CRI estimates of reliability. This was particularly clear when alpha was considered. Being alpha a reliability estimate which value also depend on the number of items (with larger scales having better alphas when other things remaining equal), it was expected that the six-item dimensions would have larger alphas than the three-item dimensions. Surprisingly, three-item dimensions with positively worded items were more reliable in terms of alpha than the six-item counterparts.

Validity results were also in line with the conclusions in Dalal and Carter (2015): negatively worded items affect validity, contrary to popular belief. The effect can even be estimated from our results. While the average correlation with criteria when relatedness, autonomy and competence were measured with

the positive items was .25, when the same dimensions were measured with the negatively worded indicators, these validity coefficients dropped to an average of .11. It was therefore clear that including negative items in the measurement instrument did affect its validity, and it did so to a large extent.

Finally, another assumption, that negatively and positively worded items measure the same construct, has also found no support in our research. On the contrary, the model that fitted the data better posited a method factor underlying the negatively worded items. Therefore, whereas a single trait or substantive dimension explained the variance in the positive items, two sources of variation (trait and method) was found to underlie the negatively worded items. Positively and negatively worded items did not measure the same construct.

As required by the not-so-good psychometric properties of the original 18-item BMPN, a proposal of using only the nine positively worded items to measure the three basic psychological needs has been made. Psychometric properties of this nine-item version of the BMPN has proved to be superior to those in the original scale in two samples of two different countries, and using to different languages, Spanish and Portuguese. Although much more research on the reliability and validity of the BMPN is needed, it seems clear that the balance of positive and negative items thwarts more than helps in achieving a good measurement. Nevertheless, it is also clear from current results that the "positive only" version of the BMPN scale with three items per dimension (the ones originally phrased positively) also lacked sufficient reliability. Therefore, our proposal is to introduce more items that reliably measure each dimension. A possibility is to rephrase the negatively worded items into positive ones. As an example the item "I felt unappreciated by one or more important people" could be rephrased as "I felt appreciated by one or more important people", and therefore, the content validity of the scale could remain unchanged.

## References

- Adie, J. W., Duda, J. L., & Ntoumanis, N. (2012). Perceived coach-autonomy support, basic need satisfaction and the well- and ill-being of elite youth soccer players: A longitudinal investigation. *Psychology of Sport and Exercise, 13*, 51-59.
- Anastasi, A. (1982). *Psychological testing* (5<sup>th</sup> Ed.). New York, NY: McMillan.
- Bachman, J. G., & O'Malley, P. M. (1986). Self-concept, self-esteem, and educational experiences: The frog pond revisited (again). *Journal of Personality and Social Psychology, 50*, 35-46.
- Bagozzi, R. P. (1993). Assessing construct validity in personality research: Applications to measures of self-esteem. *Journal of Research in Personality, 27*, 49-87.
- Brien, M., Forest, J., Mageau, G. A., Boudrias, J. S., Desrumaux, P., Brunet, L., & Morin, E. (2012). The basic psychological needs at work scale: Measurement invariance between Canada and France. *Applied Psychology: Health and Well-Being, 4*, 167-187.
- Caballo, V. E., Salazar, I. C., Arias, B., Irujo, M. J., Calderero, M., & CISO-A Research Team Spain. (2010). Validation of the social anxiety questionnaire for adults (SAQ-A30) with Spanish university students: Similarities and differences among degree subjects and regions. *Behavioral Psychology/Psicología Conductual, 18*, 5-34.
- Carmine, E. G., & Zeller, R. A. (1974). On establishing the empirical dimensionality of theoretical terms: An analytical example. *Political Methodology, 1*, 75-96.

- Carmines, E. G., & Zeller, R. A. (1979). *Reliability and validity assessment*. Beverly Hills, CA: Sage.
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing MI. *Structural Equation Modeling, 9*, 235-255.
- Corwyn, R. F. (2000). The factor structure of global self-esteem among adolescents and adults. *Journal of Research in Personality, 34*, 357-379.
- Cronbach, L. J. (1942). Studies of acquiescence as a factor in the true-false test. *Journal of Educational Psychology, 33*, 401-415.
- Dalal, D. K., & Carter, N. T. (2015). Negatively worded items negatively impact survey research. In C. E. Lance, & R. J. Vandenberg (Eds.), *More statistical and methodological myths and urban legends* (pp. 112-132). New York, NY: Routledge.
- de Haan, A. D., Soenens, B., Dekovic, M., & Prinzie, P. (2013). Effects of childhood aggression on parenting during adolescence: The role of parental psychological need satisfaction. *Journal of Clinical Child & Adolescent Psychology, 42*, 393-404.
- De Jonge, P., & Slaets, J. (2005). Response sets in self-report data and their associations with personality traits. *European Journal of Psychiatry, 19*, 209-214.
- Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York, NY: Plenum.
- Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry, 11*, 227-268.
- Deci, E. L., Ryan, R. M., Gagné, M., Leone, D. R., Usunov, J., & Kornazheva, B. P. (2001). Need satisfaction, motivation, and well-being in the work organizations of a former Eastern bloc country: A cross-cultural study of self-determination. *Personality and Psychology Bulletin, 27*, 930-942.
- DiStefano, C., & Motl, R. W. (2006). Further investigating method effects associated with negatively worded items on self-report surveys. *Structural Equation Modeling, 13*, 440-464.
- Faria, L., & Lima-Santos, N. (2012). Emotional intelligence in the Portuguese academic context: Validation studies of "The Emotional Skills and Competence Questionnaire" (ESQ). *Behavioral Psychology/Psicología Conductual, 20*, 91-102.
- Fernández-Castillo, A., Vilchez-Lara, M. J., & Sada-Lázaro, E. (2012). Satisfaction with Hospitalization Scale: Adaptation and psychometric properties. *Behavioral Psychology/Psicología Conductual, 20*, 383-400.
- Filak, V., & Sheldon, K. M. (2003). Student psychological need satisfaction and college teacher-course evaluations. *Educational Psychology, 23*, 235-247.
- Gagné, M. (2003). The role of autonomy support and autonomy orientation in prosocial behavior engagement. *Motivation and Emotion, 27*, 199-223.
- Goldsmith, R. E. (1986). Dimensionality of the Rosenberg Self-Esteem Scale. *Journal of Behavior and Personality, 1*, 253-264.
- Graham, J. M. (2006). Congeneric and essentially tau-equivalent estimates of score reliability: What they are and how to use them. *Educational and Psychological Measurement, 66*, 930-944.
- Hensley, W. E., & Roberts, M. K. (1976). Dimensions of Rosenberg's Self-Esteem Scale. *Psychological Reports, 38*, 583-584.
- Hu, L., & Bentler, P. M. (1999). Cut-off criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1-55.
- Jang, H., Kim, E.U. y Reeve, J. (2012). Longitudinal test of self-determination theory's motivation mediation model in a naturally occurring classroom context. *Journal of Educational Psychology, 104*, 1175-1188.



- Johnston, M. M., & Finney, S. J. (2010). Measuring basic needs satisfaction: Evaluating previous research and conducting new psychometric evaluations of the basic needs satisfaction in general scale. *Contemporary Educational Psychology, 35*, 280-296.
- Kaufman, P., Rasinski, K. A., Lee, R., & West, J. (1991). *National Education Longitudinal Study of 1988: Quality of the responses of eighth-grade students in NELS88*. Washington, DC: U.S. Department of Education.
- Kohn, M. L. (1977). *Class and conformity: A study of values* (2<sup>nd</sup> Ed.). Chicago, IL: University of Chicago Press.
- Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology, 140*, 5-55.
- Little, T. D. (1997). Mean and covariance structures (MACS) analyses of cross-cultural data: Practical and theoretical issues. *Multivariate Behavioral Research, 32*, 53-76.
- Marsh, H. W. (1996). Positive and negative global self-esteem: A substantively meaningful distinction or artifactors? *Journal of Personality and Social Psychology, 70*, 810-819.
- Marsh, H. W., & Grayson, D. (1995). Latent variable models of multitrait-multimethod data. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 177-198). Thousand Oaks, CA: Sage.
- Muthén, L. K., & Muthén, B. O. (2007). *Mplus User's Guide* (6<sup>th</sup> Ed.) Los Angeles, CA: Muthén & Muthén.
- Nie, Y. & Lau, S. (2009). Complementary roles of care and behavioral control in classroom management: The self-determination theory perspective. *Contemporary Educational Psychology, 34*, 185-194.
- Niemiec, C. P., Ryan, R. M., & Deci, E. L. (2009). The path taken: Consequences of attaining intrinsic and extrinsic aspirations in post-college life. *Journal of Research in Personality, 43*, 291-306.
- Nunnally, J. C. (1978). *Psychometric theory*. New York, NY: McGraw-Hill.
- Quilty, L. C., Oakman, J. M., & Risko, E. (2006). Correlates of the Rosenberg Self-Esteem Scale method effects. *Structural Equation Modeling, 13*, 99-117.
- Raufelder, D., Kittler, F., Braun, S.R., Lätsch, A., Wilkiinson, R.P., & Hoferichter, F. (2014). The interplay of perceived stress, self-determination and school engagement in adolescence. *School Psychology International, 35*, 405-420.
- Raykov, T. (2004). Behavioral scale reliability and measurement invariance evaluation using latent variable modeling. *Behavioral Therapy, 35*, 299-331.
- Reeve, J. (2012). A Self-determination Theory perspective on student engagement. In, S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 149-171). New York, NY: Springer.
- Reis, H. T., Sheldon, K. M., Gable, S. L., Roscoe, J., & Ryan, R. M. (2000). Daily well-being: The role of autonomy, competence, and relatedness. *Personality and Social Psychology Bulletin, 26*, 419-435.
- Reise, S. P. (2012). The rediscovery of bifactor measurement models. *Multivariate Behavioral Research, 47*, 667-696.
- Reise, S.P., Morizot, J., & Hays, R.D. (2007). The role of the bifactor model in resolving dimensionality issues in health outcomes measures. *Quality of Life Research, 16*, 19-31.
- Ryan, R. M. (1995). Psychological needs and the facilitation of integrative processes. *Journal of Personality, 63*, 397-427.
- Salgado, J. F., & Iglesias, M. (1995). Estructura factorial de la Escala de autoestima de Rosenberg: Un análisis factorial confirmatorio [Factorial structure of Rosenberg Self-Esteem Scale: A confirmatory factor analysis]. *Psicológica, 16*, 441-454.
- Sheldon, K. M., & Bettencourt, B. A. (2002). Psychological needs and subjective well-being in social groups. *British Journal of Social Psychology, 41*, 25-38.

- Sheldon, K. M., & Gunz, A. (2009). Psychological needs as basic motives, not just experiential requirements. *Journal of Personality, 77*, 1467-1492.
- Sheldon, K. M., & Hilpert, J. C. (2012). The Balanced Measure of Psychological Needs (BMPN) scale: An alternative domain general measure of need satisfaction. *Motivation and Emotion, 36*, 439-451.
- Sheldon, K. M., & Niemiec, C. P. (2006). It's not just the amount that counts: Balanced need satisfaction also affects well-being. *Journal of Personality and Social Psychology, 91*, 331-341.
- Sheldon, K. M., Abad, N., & Hinsch, C. (2011). A two process view of Facebook use and relatedness need-satisfaction: Disconnectedness drives use and connectedness rewards it. *Journal of Personality and Social Psychology, 100*, 766-775.
- Tomás, J. M., & Oliver, A. (1999). Rosenberg's Self-Esteem Scale: Two factors or method effects. *Structural Equation Modeling, 6*, 84-98.
- Tomás, J. M., & Oliver, A. (2004). Análisis psicométrico confirmatorio de una medida multidimensional del autoconcepto en español [Psychometric confirmatory analysis of a Spanish multidimensional measure of self-concept]. *Interamerican Journal of Psychology, 38*, 285-293.
- Tomás, J. M., Meléndez, J. C., Oliver, A., Navarro, E., & Zaragoza, G. (2010). Efectos de método en las escalas de Ryff: Un estudio en población de personas mayores [Method effects in Ryff's scales: A study on older population]. *Psicologica, 31*, 383-400.
- Tomás, J. M., Oliver, A., Galiana, L., Sancho, P., & Lila, M. (2013). Explaining method effects associated with negatively worded items in trait and state global and domain-specific self-esteem scales. *Structural Equation Modeling, 20*, 299-313.
- Vansteenkiste, M., & Ryan, R. M. (2013). On psychological growth and vulnerability: Basic psychological need satisfaction and need frustration as a unifying principle. *Journal of Psychotherapy Integration, 23*, 263-280.
- Veiga, F. H. (2013). Envolvimento dos alunos na escola: Elaboração de uma nova escala de avaliação [Students' involvement in school: Elaboration of a new assessment scale]. *International Journal of Developmental and Educational Psychology, 1*, 441-450.
- Vlachopoulos, S. P., & Michailidou, S. (2006). Development and initial validation of a measure of autonomy, competence, and relatedness in exercise: The basic psychological needs in exercise scale. *Measurement in Physical Education and Exercise Science, 103*, 179-201.
- Wang, J., Siegal, H. A., Falck, R. S., & Carlson, R. G. (2001). Factorial structure of Rosenberg's Self-Esteem scale among crack-cocaine drug users. *Structural Equation Modeling, 8*, 275-286.
- Weinstein, N., Przybylski, A. K., & Ryan, R. M. (2012). The index of autonomous functioning: Development of a scale of human autonomy. *Journal of Research in Personality, 46*, 397-413.
- Wilson, P. M., Rogers, W. T., Rodgers, W. M., & Wild, T. C. (2006). The psychological need satisfaction in exercise scale. *Exercise Psychology, 28*, 231-351.

RECEIVED: May 21, 2015

ACCEPTED: July 19, 2015

**Appendix A**

Spanish and Portuguese Positive versions of the Balanced Measure of Psychological Needs

Item number	Dimension	Item content
Spanish version		
1	Relatedness	Sentía una conexión con personas a las que les importo, y que también me importan
2	Competence	Estaba cumpliendo con éxito tareas y proyectos difíciles
3	Autonomy	Tenía la libertad de hacer las cosas a mi manera
7	Relatedness	Me sentía conectado, cercano con otras personas con otras personas importantes para mí
8	Competence	Asumí retos difíciles y los superé
9	Autonomy	Mis elecciones expresaban mi "verdadero yo"
13	Relatedness	Me sentía íntimamente unido/a a las personas con las que pasaba el tiempo
14	Competence	Manejaba bien aún las pruebas difíciles
15	Autonomy	Estaba haciendo lo que verdaderamente me interesaba
Portuguese version		
1	Relatedness	Senti-me conectado/a com pessoas as quais lhes interesse, e que também me convêm
2	Competence	Cumpri com êxito as tarefas e projetos difíceis
3	Autonomy	Tive liberdade de fazer as coisas a minha maneira
7	Relatedness	Senti-me perto e ligado a outras pessoas que são importantes para mim
8	Competence	Assumi reptos difíceis e os superei
9	Autonomy	Nas minhas eleições expressei o meu "verdadero eu"
13	Relatedness	Senti-me intimamente unido/a a pessoas com as quais passava o tempo
14	Competence	Fiz bem inclusive as coisas difíceis
15	Autonomy	Fiz o que verdadeiramente me interessava

