The Scale of Positive and Negative Experience (SPANE)

Evaluation of Measurement Invariance and Convergent and Discriminant Validity

Jovanović, V.; Lazić, M.; Gavrilov-Jerković, V.; Molenaar, D.

DOI
10.1027/1015-5759/a000540

Publication date
2020

Document Version
Final published version

Published in
European Journal of Psychological Assessment

Multistudy Report

The Scale of Positive and Negative Experience (SPANE)

Evaluation of Measurement Invariance and Convergent and Discriminant Validity

Veljko Jovanović[1], Milica Lazić[1], Vesna Gavrilov-Jerković[1], and Dylan Molenaar[2]

[1]Department of Psychology, Faculty of Philosophy, University of Novi Sad, Serbia
[2]Department of Psychology, University of Amsterdam, The Netherlands

Abstract: The Scale of Positive and Negative Experience (SPANE) is a self-report questionnaire designed to assess positive and negative emotions. In Study 1 (N = 4,250, 61.95% females, M_{age} = 28.56 years), we evaluated measurement invariance of the SPANE across gender and age using moderated factor analysis. In Study 2 (N = 200, 52.5% females, M_{age} = 21.82 years), we investigated the convergent and discriminant validity of the SPANE by examining its associations with measures of well-being and religiosity. In Study 3 (N = 160, 87.5% females, M_{age} = 20.38 years), we used a prospective design to examine associations of the SPANE with the Big Five personality traits. The results provided general support for the measurement invariance of the SPANE across age and gender, but some non-invariant items were detected as well. The analyses of latent mean differences across gender revealed that women reported higher levels of both positive and negative emotions than men, but the effect size for positive emotions was very small. Older participants reported lower levels of positive emotions and higher levels of negative emotions than younger participants. Both convergent and discriminant validity of the SPANE were supported.

Keywords: measurement invariance, validity, positive affect, negative affect, age

Positive affect (PA) and negative affect (NA) are core emotional components of subjective well-being (SWB) (Diener, Suh, Lucas, & Smith, 1999). The most widely used measure of PA and NA is the Positive and Negative Affect Schedule (PANAS; Watson, Clark, & Tellegen, 1988), which has received much criticism since its development (Thompson, 2007). The most frequently raised concerns about the PANAS have focused on the inclusion of non-emotional items and redundant items, the omission of several common emotions, and the predominance of high arousal emotions (e.g., Diener et al., 2009; Harmon-Jones, Bastian, & Harmon-Jones, 2016).

The Scale of Positive and Negative Experience (SPANE; Diener et al., 2010) has recently been developed to overcome the limitations of widely known measures of self-reported emotional experience such as the PANAS. The SPANE consists of two 6-item subscales aimed at measuring positive feelings (SPANE-P) and negative feelings (SPANE-N). Each subscale includes both general emotional terms (e.g., pleasant, unpleasant) and terms referring to specific emotions (e.g., joyful, afraid). The SPANE has several potential advantages over the PANAS, as it appears to include items with less redundancy and items with a wider range of arousal level. In addition, the SPANE includes evaluative-emotional terms (e.g., good, bad) and basic emotions (e.g., sadness, happiness), which have been found to have a high level of cross-cultural universality (e.g., Saucier, Thalmayer, & Bel-Bahar, 2014).

The use of the SPANE in well-being research has gained increasing popularity, and the scale has been translated into at least a dozen languages. The SPANE has undergone psychometric testing in several countries, including Canada (Howell & Buro, 2015), China (Li, Bai, & Wang, 2013), Germany (Rahm, Heise, & Schuldt, 2017), Italy (Giuntoli, Ceccarini, Sica, & Caudek, 2017), Japan (Sumi, 2014), Portugal (Silva & Caetano, 2013), and South Africa (Du Plessis & Guse, 2017). The correlations between the SPANE-P and positive indicators of SWB, such as happiness, PA (measured by the PANAS), and life satisfaction have typically been in the range from .60 to .70, whereas the correlations between the SPANE-N and NA (measured by the PANAS) have been in the range from .70 to .77 (e.g., Giuntoli et al., 2017; Rahm et al., 2017). Adequate internal consistency indices of the SPANE subscales, ranging from .79 to .92, were found in previous studies (e.g., Du Plessis & Guse, 2017; Howell & Buro, 2015; Li et al., 2013; Silva & Caetano, 2013). In addition, one-month test-retest reliabilities of the SPANE-P and the SPANE-N in the German validation study were .62 and .64, respectively (e.g., Rahm et al., 2017). However, despite its prevalent use

in well-being research, the discriminant validity of the SPANE has not yet been sufficiently evaluated.

Studies investigating the structural validity of the SPANE have consistently found support for the two correlated factor models (e.g., Giuntoli et al., 2017; Rahm et al., 2017; Sumi, 2014), but only a few studies have evaluated measurement invariance (MI) of the SPANE. MI testing provides a robust test on whether the scale measures the same constructs in the same way across different groups (Sass, 2011). Therefore, the evaluation of MI is crucial in studies that focus on the data comparison across groups. Giuntoli and colleagues (2017) found that an Italian version obtained strict MI across administration methods (paper-and-pencil and Internet) and scalar MI across different groups (unemployed individuals seeking work and a healthy control group), and Li and colleagues (2013) reported strict MI of a Chinese version across gender and age.

Age and gender differences in PA and NA have been widely studied in the field of SWB (e.g., Lucas & Gohm, 2000; Steptoe, Deaton, & Stone, 2015). In addition, the question of common and unique correlates of PA and NA among men and women (e.g., Isaacowitz & Smith, 2003) and among people in different periods of life (e.g., Masumoto, Taishi, & Shiozaki, 2016) has also puzzled many researchers. However, comparisons of both mean levels and correlates of PA and NA across different groups have typically been conducted under an untested assumption about MI. Without the establishment of MI, comparisons of group means or cross-group relationships are likely to be invalid and meaningless (Vandenberg & Lance, 2000).

The Present Research

The goal of Study 1 was twofold: (1) to evaluate the structural validity of the SPANE; (2) to examine the MI of the SPANE across gender and age groups, and to test for latent mean differences across these variables. Study 2 examined the convergent and discriminant validity of the SPANE by examining its associations with the PANAS, convergent measures of well-being (life satisfaction, depression, the presence of meaning in life), and two discriminant measures (the search for meaning in life and frequency of religious attendance). Previous studies have shown that PA and NA are closely related to life satisfaction (e.g., Kuppens, Realo, & Diener, 2008) and the presence of meaning in life (e.g., Boyraz, Lightsey, & Can, 2013). In addition, depression has consistently been found in previous studies to have a strong positive correlation with NA and a strong negative correlation with PA (e.g., Henry & Crawford, 2005). The search for meaning in life and religiosity were used to evaluate the SPANE's discriminant validity, since previous studies have typically found nonsignificant or low correlations between these constructs and PA and NA (e.g., Diener, Tay, & Myers, 2011; Park, Park, & Peterson, 2010). Study 3 used a prospective design to examine the associations between personality traits and PA and NA. The main goal of this study was to further evaluate the convergent validity of the SPANE by investigating the predictive value of personality traits in predicting PA and NA 1 year later. We used both the SPANE and the PANAS in these analyses in order to compare the predictive value of personality traits in predicting PA and NA using two alternative measures of affect. Previous studies have shown that personality traits are among the most robust predictors of PA and NA. More specifically, extraversion and neuroticism have been shown to be the strongest predictors of PA and NA, respectively (Steel, Schmidt, & Shultz, 2008). Agreeableness and conscientiousness have also been found to predict PA and NA (Soto, 2015), whereas the contribution of openness to experience has typically been either small or negligible (Soto, 2015; Steel et al., 2008).

Study 1

Methods

Participants and Procedure

A total of 4,250 Serbian participants (61.95% females) were recruited for the present study. The mean age was 28.56 years (SD = 12.60, range: 15–85 years). In establishing the factor structure (see Factor Structure section below), five age bands were used: 15–18 years (N = 768; 71.06% females), 19–24 years (N = 1,568; 61.15% females), 25–34 years (N = 848; 63.03% females), 35–49 years (N = 590; 58.22% females), and older than 50 years (N = 449; 51.79% females). The age groups were created to comprise different developmental periods and to obtain large sample sizes.

Participation in the study was anonymous and voluntary, and the participants did not receive any compensation for taking part in it. The questionnaire was administered in a standard paper-and-pencil format. Using convenience sampling, participants aged 15–18 years were recruited from three state schools in Serbia, whereas participants aged 19–24 years were recruited from the University of Novi Sad. Participants aged 25 years and older were recruited through a snowball sampling method, in which undergraduate students at the University of Novi Sad contacted friends, acquaintances, and members of their family to complete the questionnaire.

Instruments

Scale of Positive and Negative Experience (SPANE; Diener et al., 2010) is a 12-item scale, with 6 items designed to

assess positive emotional experiences (SPANE-P), and 6 items to assess negative emotional experiences (SPANE-N). Each item is rated on a 5-point scale (1 = *very rarely or never*, 5 = *very often or always*), and rating are made over a 4-week time period. A Serbian version of the SPANE was used in the present study, which demonstrated adequate reliability, structural and incremental validity in previous studies (Jovanović, 2015). The Serbian version of the SPANE has been developed using a back-translation procedure and following a standard guidelines for the development and translation of self-report questionnaires (for details, see Tran, 2009). In the present study, both the SPANE-P (α = .90) and the SPANE-N (α = .84) demonstrated good internal consistency. Means, standard deviations, skewness, kurtosis, and internal consistency of the SPANE-P and the SPANE-N, broken down by gender and age groups are shown in Electronic Supplementary Material (ESM 1).

Data Analysis

Missing Data

In the full sample of 4,250 participants, 4,179 participants had a full data record, 33 participants had missing values on a single item, 11 participants had missing on 2 or more items, and for 27 participants all item scores were missing. In the analysis below, these 27 participants with a fully missing record were omitted from the analysis. All other participants (with at least one observation) were included in the analyses.

Factor Structure

To establish the factor structure, we used Confirmatory Factor Analysis (CFA) in Lavaan (Rosseel, 2012). We wanted to establish (1) that the theoretically assumed two-factor model fits acceptable to the data, and (2) that the two-factor model fits better than the competing one-factor model suggesting that we can indeed separate the two affect factors. The CFA is conducted within the total sample, the male sample and the female sample, and within each age group. All items were explicitly treated as categorical and the parameters were estimated using weighted least squares estimation.[1] Several fit indices were used to evaluate the model: The Root Mean Square Error of Approximation (RMSEA), Comparative Fit Index (CFI), and the Standardized Root Mean Squared Residual (SRMR). Good model fit was defined by the following criteria: SRMR value less than .08, RMSEA values from .06 or less, and CFI above .95, whereas values between .90 and .95 for CFI, and between .06 and .08 for RMSEA were considered acceptable (Hu & Bentler, 1999).

Measurement Invariance

As we wanted to establish MI with respect to a continuous variable (age), we used the method of moderated factor analysis (MFA) (Bauer, 2017; Molenaar, Dolan, Wicherts, & van der Maas, 2010). Advantage of MFA over the more traditional multi-group CFA is that (1) the age variable does not need to be categorized so that its continuous nature can be retained in the MI analysis, and (2) MI can be tested more easily with respect to gender and age simultaneously such that possible interactions between age and gender in MI can be accounted for (e.g., factor loadings may be increasing across age, but with a different rate for males as compared to females). In addition, if MI is tested with respect to age and gender separately, one has to assume MI for the variable not under consideration (e.g., if MI is tested with respect to age only, one assumes MI with respect to gender in each age group; similarly, if MI is tested with respect to gender only, one assumes MI with respect to age in the male and female sample). Note that MFA with a categorical moderator (gender in this case) is equivalent to the traditional multi-group CFA approach (see Bauer & Hussong, 2009). See ESM 2 for a detailed explanation of the procedure followed.

Because the likelihood ratio test is relatively sensitive to sample size, we consider the following fit indices to evaluate which of the candidate models is the best fitting model: the Bayesian Information Criterion (BIC), Akaike's Information Criterion (AIC), the sample size adjusted BIC (saBIC), and the Deviance Information Criterion (DIC). For all fit indices it holds that a better fitting model has a smaller value on the respective fit index. The SPANE-P and the SPANE-N scales are analyzed separately to not complicate the models too much.[2] In addition, all items are explicitly treated as categorical. Models were fit in Mplus (Muthén & Muthén, 2015) using marginal maximum likelihood estimation with robust standard errors.

Results

Factor Structure of the SPANE

As can be seen in ESM 3, the one-factor model fitted poor in terms of the RMSEA (values between .102 and .154) and SRMR (values between .075 and .110), and good in terms of

[1] Note that in CFA, pragmatically, one can assume categorical variables as approximately continuous if the variable has five or more response categories with a symmetrical distribution (Rhemtulla, Brosseau-Liard, & Savalei, 2012). However, although we have data with five response categories, the present data distribution departs from a symmetrical distribution due to the disproportional use of the upper or lower categories (depending on the question). This is common to affect research where you hardly observe symmetrical distributions for the observed data.

[2] Testing moderation of NA and PA simultaneously in a two-factor model requires a moderated correlation between the factors. Although this is possible in principle model complexity increases drastically (as also noted by Bauer, 2017).

the CFI (values between .963 and .986). The two-factor model fitted better according to the RMSEA (values between .026 and .061), the SRMR (values between .032 and .042), and the CFI (values between .997 and .998). We additionally conducted an Exploratory Factor Analysis (EFA; results available upon request) to see whether there are any violations of the factor structure as specified in the theoretically motivated two-factor model. See ESM 4 for the standardized factor loadings in the EFA and the CFA. As can be seen, there are no important departures from the theoretical structure. Therefore, we accept the confirmatory two-factor model.

Measurement Invariance

SPANE-P

See Table 1 for the results of the SPANE-P scale. As discussed in ESM 2, we started with a moderation model in which all loadings and thresholds are moderated by all moderators (Model#1). Next, we dropped the interaction moderators (Gender × Age and Gender × Age-Squared; G×A and G×A2 in Table 1) from the loadings and added these moderators as moderators of the factor variance (Model#2). As can be seen from Table 1, all fit indices indicated that this model fitted better than the previous model. Next, we dropped the age moderators (age and age-squared; A and A2 in Table 1) and added these moderators as moderators of the factor variance (Model#3). Again this model fitted better than the previous model according to all fit indices. Then, we dropped the gender moderator (G in Table 1) and added these moderators as moderators of the factor variance (Model#4). Model fit was again better than the models before. Next we started considering the thresholds. First, we dropped the interaction moderators from the thresholds and added these moderators as moderators of the factor means (Model#5). Model#5 fitted better than the models before. Then, we dropped the age moderators and added these moderators as moderators of the factor mean (Model#6). As shown in Table 1, the AIC, the saBIC, and the DIC indicated that model fit became worse (Model#6 vs. Model#5). We therefore consulted the source of misfit in Model#6 by adding moderation of the thresholds in the presence of moderation of the factor mean (we used the first item as anchor). It appeared that for item 12 ("contented"), there was a direct effect of age on the thresholds in addition to the effect via the factor mean. Therefore, we slightly revised Model#6 and added moderation of the thresholds of item 12 by age. The resulting model (Model#6') fitted better than the model from the previous step (Model#5, i.e., the model with full moderation of the thresholds by age and age-squared) according to all fit indices but AIC. However, as all other fit indices indicated that Model#6' was the better fitting model, we accepted this model. In this model, the path coefficient of the path from age directly to item 12 was estimated to be 0.224 (*SE* = .035) indicating that older participants score higher on this item than one would expect based on their latent variable score. The AIC is known to be overly strict to parsimony. In the final model, Model#7, we dropped moderation of the thresholds by gender and added moderation of the factor mean by gender. This model fitted better as compared to Model#6' according to all fit indices but AIC. We therefore accept this model as the final model and conclude that the SPANE-P scale is measurement invariant with respect to gender and age, with the exception that item 12 is not fully measurement invariant with respect to age.

SPANE-N

We followed a similar procedure for the SPANE-N scale. As shown in Table 2, all models improved model fit up until the full-MI model (Model#7). This model fitted worse as compared to Model#6 according to all fit indices. We consulted the item specific moderation effects of the thresholds and found items 4 ("bad"), 8 ("sad"), and 9 ("afraid") to show item specific moderation effects. We therefore revised Model#7 and added these effects to the thresholds. The three paths were estimated to be equal to 0.378 (*SE* = .091), 0.630 (*SE* = .077), and 0.679 (*SE* = .069) for items 4,

Table 1. Fit indices of the different moderation models fit to the data from the SPANE-P scale

	Moderators on parameters				Fit indices			
Model	Loadings	Thresholds	Factor mean	Factor variance	AIC	BIC	saBIC	DIC
1 (No-MI)	G, A, A2, G×A, G×A2	G, A, A2, G×A, G×A2			50,779	51,349	51,063	51,184
2	G, A, A2	G, A, A2, G×A, G×A2		G×A, G×A2	48,324	48,831	48,577	48,684
3	G	G, A, A2, G×A, G×A2		A, A2, G×A, G×A2	48,237	48,681	48,458	48,552
4		G, A, A2, G×A, G×A2		G, A, A2, G×A, G×A2	48,233	48,645	48,438	48,525
5		G, A, A2	G×A, G×A2	G, A, A2, G×A, G×A2	**48,220**	48,568	48,394	48,467
6		G	A, A2, G×A, G×A2	G, A, A2, G×A, G×A2	48,266	48,552	48,409	48,469
6'		G, A [item 12]	A, A2, G×A, G×A2	G, A, A2, G×A, G×A2	48,231	48,522	48,376	48,438
7 (MI)		A [item 12]	G, A, A2, G×A, G×A2	G, A, A2, G×A, G×A2	48,235	**48,495**	**48,365**	**48,420**

Note. Best values for the fit indices are in boldface; MI = measurement invariance, G = gender, A = age, A2 = age-squared, G×A = gender × age, G×A2 = gender × age-squared.

Table 2. Fit indices of the different moderation models fit to the data from the SPANE-N scale

	Moderators on parameters				Fit indices			
Model	Loadings	Thresholds	Factor mean	Factor variance	AIC	BIC	saBIC	DIC
1 (No-MI)	G, A, A2, G×A, G×A2	G, A, A2, G×A, G×A2			59,153	59,723	59,437	59,558
2	G, A, A2	G, A, A2, G×A, G×A2		G×A, G×A2	57,133	57,640	57,386	57,493
3	G	G, A, A2, G×A, G×A2		A, A2, G×A, G×A2	57,072	57,515	57,293	57,387
4		G, A, A2, G×A, G×A2		G, A, A2, G×A, G×A2	57,070	57,482	57,275	57,362
5		G, A, A2	G×A, G×A2	G, A, A2, G×A, G×A2	57,071	57,419	57,244	57,318
6		G	A, A2, G×A, G×A2	G, A, A2, G×A, G×A2	**57,104**	57,390	**57,247**	**57,307**
7 (MI)		G, A [item 12]	A, A2, G×A, G×A2	G, A, A2, G×A, G×A2	57,255	57,509	57,382	57,435
7' (MI)		G [item 4, 8, 9]	G, A, A2, G×A, G×A2	G, A, A2, G×A, G×A2	57,116	**57,389**	57,252	57,310

Note. Best values for the fit indices are in boldface. MI = measurement invariance; G = gender; A = age; A2 = age-squared; G×A = Gender × Age; G×A2 = gender × age-squared.

8, and 9, respectively. These estimates indicate that for these items, females score higher than one would expect based on their latent variable scores. The resulting Model#7' fitted better according to BIC, but worse according to AIC, saBIC, and DIC. However, the differences for the saBIC and DIC are very small, and the AIC is known to be conservative toward restrictive models. We therefore conducted a traditional multi-group CFA in which we tested the loadings and thresholds to be invariant across males and females. According to the results, the loadings can be considered equal across gender as the RMSEA and CFI improves when we fixed the loadings to be equal (metric model). However, when we fixed the thresholds (scalar model) the RMSEA and CFI deteriorated. As judged by the modification indices, the most obvious source of misfit is item 9. We therefore freed the thresholds for item 9 across groups and model fit improved, however, the resulting model still fitted worse as compared to the metric model. Again consulting the modification indices showed that the largest misfit was in the threshold of item 8. After freeing, these thresholds across groups resulted in a model that fitted better than the metric model.

Combining the results from the MFA and the traditional multi-group CFA, we conclude that generally, the SPANE-N scale is measurement invariant with respect to age, and partially measurement invariant with respect to gender as items 8 and 9 violate measurement invariance. Results concerning item 4 are unclear as the final moderated factor models are hard to distinguish. In addition, item 4 is not flagged in the traditional analysis.

Latent Mean Differences

If we focus on the results of the final models (Model#7) for both the SPANE-P and the SPANE-N, we can conclude the following: There was a main effect for gender with females reporting higher levels of both PA ($z = 1.977$, $p < .05$) and NA ($z = 5.255$, $p < .001$) than males (reference group). Although significant, the effect size for PA was very small.

For PA, there was a linear effect of age ($z = -9.080$, $p < .001$), with older participants reporting lower levels of PA. A quadratic effect of age was also significant ($z = 2.302$, $p < .05$), but the effect size was small. For PA, there was no interaction between age and gender. For NA, there was a linear effect of age ($z = 5.057$, $p < .001$), with older participants reporting higher levels of NA. No quadratic effect of age was found for NA. For NA there was also an interaction between gender and age ($z = -2.216$, $p < .05$) with the older females reporting lower NA than the older males, and the younger females reporting higher NA than the younger males. See ESM 5 for a graphical representation of these results.

Study 2

Methods

Participants and Procedure

A total of 200 undergraduate students (52.5% females) from the University of Novi Sad were recruited for Study 2. The mean age was 21.82 years ($SD = 2.48$, range: 19–34 years). The majority of participants were unemployed (84.5%), were involved in a romantic relationship (57.5%), and declared themselves as Orthodox Christians (62.5%). The participants rated their financial situation as poor (7%), only fair (26.7%), good (39.7%), very good (22.6%), and excellent (4%). All participants completed an informed consent. The questionnaires were administered by a research assistant in a group setting using a standard paper-and-pencil format.

Instruments

The SPANE as described in Study 1 was also used in Study 2. Internal consistencies (Cronbach's α) for the present study of the SPANE-P and the SPANE-N were .91 and .83, respectively. In addition to the SPANE, we used the following instruments in Study 2:

The Positive and Negative Affect Schedule (PANAS; Watson et al., 1988) is a 20-item scale designed to assess PA (PANAS-PA) and NA (PANAS-NA), using a 5-point scale from 1 (= *very slightly or not at all*) to 5 (= *extremely*). We used the Serbian translation of the PANAS which demonstrated adequate psychometric properties in previous studies (Jovanović & Gavrilov-Jerković, 2016). Participants were asked to report how they felt during the last month. Both subscales obtained a Cronbach's α value of .83 in the present study.

Satisfaction with Life Scale (SWLS; Diener, Emmons, Larsen, & Griffin, 1985) consists of 5 items (e.g., I am satisfied with my life). It uses a 7-point Likert type scale (1 = *strongly disagree* to 7 = *strongly agree*). A Serbian version of the SWLS has been found to have adequate reliability and validity (Vasić, Šarčević, & Trogrlić, 2011), and it has demonstrated MI across age, gender, and time (Jovanović, 2019). The Cronbach's α of the scale was .84 in the present sample.

Depression subscale of the Depression Anxiety and Stress Scale (DASS-21; Lovibond & Lovibond, 1995) was used to measure symptoms of depression. This subscale includes 7 items (e.g., I felt down-hearted and blue), rated on a 4-point scale, ranging from 0 (= *did not apply to me at all*) to 3 (= *applied to me very much, or most of the time*). A Serbian version of the DASS-21 demonstrated strong reliability and validity in previous studies (Jovanović, Gavrilov-Jerković, Žuljević, & Brdarić, 2014). Internal consistency was adequate (α = .84) in this study.

Meaning in Life Questionnaire (MLQ; Steger, Frazier, Oishi, & Kaler, 2006) consists of two 5-item subscales: Presence of Meaning (e.g., "My life has a clear sense of purpose") and Search for Meaning (e.g., "I am seeking a purpose or mission for my life"). Items are rated on a 7-point scale (1 = *absolutely untrue*, 7 = *absolutely true*). In this study, Cronbach's α values were .87 and .85 for the Presence and the Search subscales, respectively.

Frequency of religious attendance was assessed using a question "How often do you attend a place of worship (e.g., church, mosque, synagogue) for religious reasons?", rated on a 8-point scale from 1 (= *never*) to 8 (= *almost every day or every day*).

Data Analysis

To establish the convergent and discriminant validity of the SPANE, we specified a structural equation model including the items and their underlying factors as described above (SPANE-P, SPANE-N, PANAS-PA, PANAS-NA, Life satisfaction, Depression, the Presence of meaning, and the Search for meaning). In addition, the variable "Frequency of Religious Attendance" is added to the model. Next, the correlations of the SPANE-P and the SPANE-N factors with the well-being factors and the Religiosity variable are examined. A recently recommended guidelines by Gignac and Szodorai (2016) for interpreting the magnitude of correlations were used in the present study (.15 = small, .25 = medium, and .35 = large). These guidelines were developed through the quantitative investigation of a total of 708 observed correlations derived from the sample of 87 meta-analyses (for details, see Gignac & Szodorai, 2016). All observed variables are explicitly treated as categorical, and parameter estimation is conducted by means of weighted least squares estimation in Mplus (Muthén & Muthén, 2015).

Results

Convergent and Discriminant Validity of the SPANE

As shown in Table 3, the SPANE-P had high positive correlations with positive indicators of well-being and high negative correlations with the Depression factor and the PANAS-NA factor. The SPANE-N factor had the strongest correlations with the PANAS-NA factor and the Depression factor, and moderate to large negative correlations with positive indicators of well-being. In comparison with the corresponding PANAS factors, the SPANE-N and SPANE-P consistently yielded slightly higher correlations with convergent measures of well-being. Testing for differences between these correlations revealed that the SPANE-N in comparison to the PANAS-NA had significantly stronger correlations with Life Satisfaction (z = 2.00, p < .05) and with Depression (z = 2.57, p < .01) but not with the Presence of Meaning factor. In addition, the SPANE-P in comparison to the PANAS-PA had significantly stronger correlations with Life Satisfaction (z = 2.00, p < .05) but not with Depression or the Presence of Meaning. The discriminant validity of both the SPANE-P and the SPANE-N was supported by small correlations with the Religiosity variable and the Search for Meaning factor.

Study 3

Methods

Participants and Procedure

At Time 1, a total of 411 undergraduate students (84% females; M_{age} = 19.99 years, SD = 1.36, range: 18–27 years) at the University of Novi Sad were recruited for the present study. The final sample comprised 160 students (87.5% females; M_{age} = 20.38 years, SD = 0.96, range: 19–25 years) who completed measures at both Time 1 and Time 2 (after 1 year). Participants completed a measure of personality traits at Time 1, whereas at Time 2 they completed the two measures of PA and NA. The questionnaires were administered in a group setting (during class time) through

Table 3. Convergent and discriminant validity coefficients (standard error) of the SPANE

	1	2	3	4	5	6	7	8
1. SPANE-P	–							
2. SPANE-N	−.74 (.04)**	–						
3. PANAS-PA	.81 (.03)**	−.56 (.05)**	–					
4. PANAS-NA	−.63 (.05)**	.85 (.02)**	−.56 (.05)**	–				
5. Life satisfaction	.65 (.05)**	−.55 (.05)**	.55 (.05)**	−.45 (.05)**	–			
6. Depression	−.66 (.04)**	.68 (.03)**	−.58 (.05)**	.59 (.04)**	−.59 (.04)**	–		
7. Presence of meaning	.48 (.06)**	−.36 (.06)**	.42 (.07)**	−.31 (.06)**	.61 (.05)**	−.43 (.06)*	–	
8. Search for meaning	.00 (.07)	.07 (.07)	.08 (.07)	.02 (.07)	.01 (.07)	.13 (.07)	−.09 (.06)	–
9. Religiosity	.18 (.07)**	−.09 (.07)	.16 (.07)*	−.06 (.07)	.17 (.08)*	−.13 (.07)*	.28 (.07)**	.03 (.07)

Note. $^{**}p < .01$, $^{*}p < .05$.

a paper-and-pencil format. All participants completed an informed consent. Participation in a study was voluntary and participants received no compensation for participating in the study.

Instruments

The SPANE and the PANAS (as described in Study 2) were used to assess PA and NA in Study 3. In the present sample, Cronbach's α values for SPANE-P, SPANE-N, PANAS-PA, PANAS-NA were .89, .83, .86, and .86, respectively.

In addition to these two measures, we used the Big Five Inventory (BFI; John & Srivastava, 1999) to measure personality. The BFI consists of 44 items designed to assess five dimensions of personality: Extraversion (α = .81), Neuroticism (α = .82), Openness to experience (α = .87), Agreeableness (α = .75), and Conscientiousness (α = .84). Items were rated on a 5-point Likert-type scale, ranging from 1 (= *strongly disagree*) to 5 (= *strongly agree*). A Serbian translation of the BFI showed good psychometric properties in previous studies (e.g., Čolović, Smederevac, & Mitrović, 2014).

Data Analysis

A structural equation model is specified in which the two SPANE, the two PANAS and the five BFI factors are operationalized in terms of their corresponding items. Next, the five BFI factors at Time 1 are used as predictors, whereas PA and NA as measured by the SPANE and the PANAS at Time 2 are used as criteria. In addition, age and gender are added to the regression as control variables. Note that the resulting analysis is a multivariate (latent) regression model in which the criterion variables are modeled simultaneously in terms of the predictor and control variables. As a result, the correlation between the criterion variables is explicitly accounted for. All variables are explicitly treated as categorical and parameters are estimated using weighted least squares estimation in Mplus (Muthén & Muthén, 2015). See ESM 6 for the correlations between the factors.

Results

Personality Traits Predicting PA and NA 1 Year Later

As shown in Table 4, Extraversion was a significant predictor of both SPANE-P (β = .43, $p < .01$) and PANAS-PA (β = .61, $p < .01$). Neuroticism and Conscientiousness were significant predictors of both SPANE-N (β's = .44, −.24, respectively, p's < .01) and PANAS-NA (β's = .48, −.24, respectively, p's < .01). Significant predictors of PANAS-PA were also Agreeableness (β = −.23, $p < .01$), Conscientiousness (β = .19, $p < .05$), and Openness (β = .14, $p < .05$).

Discussion

The results of the present research support the two-factor structure of the SPANE, which is in line with previous studies (Giuntoli et al., 2017; Rahm et al., 2017). The results of moderated factor analyses provided general support for the SPANE's invariance across age and gender. However, there was also evidence of non-invariant items. The SPANE-P item "contented" was found to violate measurement invariance with respect to age, with the item scores of older people being higher than one would expect based on their factor scores. In addition, the "sad" and "afraid" items in the SPANE-N were found to violate measurement invariance with respect to gender, with the item scores from females being higher than one would expect based on their factor scores. The non-invariance of the two NA items across gender is in accordance with widely reported gender differences in experience and expression of internalizing negative emotions such as sadness and fear (Chaplin, 2015). The non-invariance of the "contented" item across age indicated that older individuals ascribe different meaning to this emotion term, which captures a low-arousal emotion that creates the urge to savor current life conditions and recent experiences (Fredrickson, 1998). It has been shown that older adults prefer low arousal positive

Table 4. Predictive value of personality traits for positive and negative affect

	SPANE-P		SPANE-N		PANAS-PA		PANAS-NA	
	β	*SE*	β	*SE*	β	*SE*	β	*SE*
Extraversion	.43**	.10	−.05	.11	.61**	.10	.07	.10
Neuroticism	−.15	.09	.44**	.09	−.14	.08	.48**	.08
Agreeableness	−.14	.09	−.03	.09	−.23**	.08	−.05	.09
Conscientiousness	.16	.08	−.24**	.07	.19*	.08	−.24**	.08
Openness	−.07	.07	.03	.07	.14*	.06	.06	.07
Age	.13	.09	−.15	.09	.15	.09	−.14	.08
Gender	−.02	.09	.08	.09	−.12	.08	.15*	.07
R^2	.30	.06	.37	.06	.41	.08	.34	.06

Note. $^{**}p < .01$, $^{*}p < .05$.

emotions in their daily life (Bjalkebring, Västfjäll, & Johansson, 2015), which might explain the different meanings attached to this item among older and younger individuals. Our findings are in line with the results of Du Plessis and Guse (2017), who identified "contented" and "afraid" as two problematic items using the Rasch analysis. In sum, our results suggest that meaningful comparisons across gender and age can be made using the SPANE, but the results need to be interpreted with some caution because three non-invariant items were identified.

The correlations between the SPANE-P and the SPANE-N factors were consistently around −.70 across different groups, which is comparable to the correlations found in previous studies (e.g., Giuntoli et al., 2017; Li et al., 2013; Rahm et al., 2017). It is important to note that these correlations are much stronger than those typically found between the PANAS-PA and the PANAS-NA (e.g., Crawford & Henry, 2004; Rush & Hofer, 2014). This is as expected since the SPANE includes several pairs of opposite items (e.g., pleasant/unpleasant, happy/sad) which increases an inverse relationship between its subscales. On the other hand, the PANAS was explicitly designed to provide independent measures of PA and NA since its development was based on the model of affect that postulates two orthogonal dimensions. A high negative correlation between the SPANE subscales calls into question the conceptual distinction between positive and negative experiences as measured by the SPANE. However, the diverging pattern of observed cross-sectional associations between the SPANE subscales and other variables indicate that the SPANE-P and the SPANE-N reflect highly related, yet distinguishable constructs. For example, neuroticism, agreeableness, and conscientiousness have been shown to be more closely associated with the SPANE-N than the SPANE-P, whereas extraversion, life satisfaction, and the presence of meaning have demonstrated stronger associations with the SPANE-P than the SPANE-N. These findings suggest that the two SPANE subscales have somewhat distinct relationships to indicators of well-being and personality traits.

It is important to note that the SPANE and the PANAS subscales showed a highly similar pattern of associations with other variables. This indicates that the two scales exhibit comparable convergent validity and generally perform similarly regarding relationships with convergent measures. Moreover, the same-valenced SPANE and PANAS scales (i.e., SPANE-P/PANAS-PA, SPANE-N/PANAS-NA) had an average correlation of .86 across the two studies, whereas the opposite-valenced scales (i.e., SPANE-P/PANAS-NA, SPANE-N/PANAS-PA) had a substantially lower average correlation of −.58.

Our results support both convergent and discriminant validity of the SPANE. The correlations between the SPANE and measures of well-being found in our sample are comparable to those reported in previous studies (e.g., Giuntoli et al., 2017; Rahm et al., 2017). Both SPANE subscales evidenced strong discriminant validity as indicated by low or nonsignificant correlations with the search for meaning and religiosity. The pattern of prospective associations between personality traits and the SPANE subscales provided additional support for this scale's convergent validity, and further supported the distinction between the two SPANE subscales. In accordance with previous studies (e.g., Soto, 2015), extraversion was a significant predictor of the SPANE-P, whereas neuroticism and conscientiousness were significant predictors of the SPANE-N. The fact that extraversion served as a predictor only for the SPANE-P but not for the SPANE-N, whereas the opposite was true for neuroticism and conscientiousness, further supports the argument that the SPANE-P and the SPANE-N should not be treated as the opposite poles of a single affective dimension. When PANAS subscales were used as outcome measures in these analyses, the same pattern of results was obtained for the PANAS-NA and the SPANE-N, whereas some differences were found when the SPANE-P and the PANAS-PA were compared. Most notably, extraversion was more strongly associated with the PANAS-PA than with the SPANE-P, which clearly indicates a greater semantic overlap between the PANAS-PA and extraversion items.

The finding that females reported higher levels of both PA and NA than males is only partially in accordance with previous research. Women consistently report higher NA than men in studies using both the SPANE (e.g., Li et al., 2013) and other measures of unpleasant emotions (e.g., Lucas & Gohm, 2000). However, the findings on gender differences in PA are mixed and inconsistent (for a review, see Batz & Tay, 2018), suggesting that gender differences in PA may be moderated by various factors (e.g., culture and age) or may depend on the scale used to assess PA.

We found age had a clear effect on both NA and PA. Older adults reported both lower levels of positive emotions and higher levels of negative emotions than younger participants. These results are in line with the findings of Steptoe and colleagues (2015), who found that, contrary to the usual U-shaped relationship between age and well-being, individuals from post-communist countries show a progressive decline in well-being with age. The finding that the interaction between gender and age was significant only for NA but not for PA provides further support that despite high negative correlations between the SPANE-P and the SPANE-N, these two subscales capture distinctive emotional experiences and not mere opposites.

Limitations and Future Directions

It is important to note several limitations of the present study. First, the study's cross-sectional design did not allow for testing MI over time. Future studies should evaluate the validity of the SPANE scores using a longitudinal design. Second, a single cultural context precludes the assessment of cross-cultural invariance of the SPANE. To our knowledge, no study has yet investigated the cross-cultural utility of the SPANE. Since high-arousal emotions have been found to be more valued and more frequently experienced in Western than in Eastern cultures (Lim, 2016), cross-cultural research of emotions and well-being would benefit from investigating whether the scores obtained using the SPANE have higher levels of comparability across cultures than the PANAS. Third, the present study did not focus on the direct comparison of the psychometric performance of the SPANE and the PANAS, so future studies should carefully compare different self-report measures of affect, since the present findings suggest some subtle differences between the two scales. Fourth, samples used in the present study to evaluate the SPANE's convergent and discriminant validity included only undergraduate students, which limits the generalizability of our findings. Finally, the choice of measures for testing convergent and discriminant validity was not comprehensive. Therefore, further research on the SPANE's validity is warranted.

Electronic Supplementary Materials

The electronic supplementary materials are available with the online version of the article at https://doi.org/10.1027/1015-5759/a000540

ESM 1. Descriptive statistics and internal consistency for SPANE subscales

ESM 2. Explanation of the exact procedure of testing for MI using moderated factor analysis

ESM 3. Fit indices for structural models of the SPANE using CFA

ESM 4. Standardized factor loadings of the two-factor solution obtained using CFA and of the two-factor solution obtained using EFA

ESM 5. Graphical representation of the latent mean differences across age for females and males for PA and NA

ESM 6. Correlations among the factors in the latent regression of Study 3

References

Batz, C., & Tay, L. (2018). Gender differences in subjective well-being. In E. Diener, S. Oishi, & L. Tay (Eds.), *Handbook of well-being*. Salt Lake City, UT: DEF.

Bauer, D. J. (2017). A more general model for testing measurement invariance and differential item functioning. *Psychological Methods, 22*, 507–526. https://doi.org/10.1037/met0000077

Bauer, D. J., & Hussong, A. M. (2009). Psychometric approaches for developing commensurate measures across independent studies: Traditional and new models. *Psychological Methods, 14*, 101–125. https://doi.org/10.1037/a0015583

Bjalkebring, P., Västfjäll, D., & Johansson, B. E. A. (2015). Happiness and arousal: Framing happiness as arousing results in lower happiness ratings for older adults. *Frontiers in Psychology, 6*, 706. https://doi.org/10.3389/fpsyg.2015.00706

Boyraz, G., Lightsey, O. R., & Can, A. (2013). The Turkish version of the meaning in life questionnaire: Assessing the measurement invariance across Turkish and American adult samples. *Journal of Personality Assessment, 95*, 423–431. https://doi.org/10.1080/00223891.2013.765882

Chaplin, T. M. (2015). Gender and emotion expression: A developmental contextual perspective. *Emotion Review, 7*, 14–21. https://doi.org/10.1177/1754073914544408

Crawford, J. R., & Henry, J. D. (2004). The Positive and Negative Affect Schedule (PANAS): Construct validity, measurement properties and normative data in a large non-clinical sample. *British Journal of Clinical Psychology, 43*, 245–265. https://doi.org/10.1348/0144665031752934

Čolović, P., Smederevac, S., & Mitrović, D. (2014). Velikih pet plus dva: Validacija skraćene verzije [Big Five Plus Two: Validation of a short version]. *Primenjena Psihologija, 7*, 227–254.

Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The Satisfaction with Life Scale. *Journal of Personality Assessment, 49*, 71–75. https://doi.org/10.1207/s15327752jpa4901_13

Diener, E., Suh, E. M., Lucas, R. E., & Smith, H. L. (1999). Subjective well-being: Three decades of progress. *Psychological Bulletin, 125*, 276–302. https://doi.org/10.1037/0033-2909.125.2.276

Diener, E., Tay, L., & Myers, D. G. (2011). The religion paradox: If religion makes people happy, why are so many dropping out? *Journal of Personality and Social Psychology, 101*, 1278–1290. https://doi.org/10.1037/a0024402

Diener, E., Wirtz, D., Biswas-Diener, R., Tov, W., Kim-Prieto, C., Choi, D.-W., & Oishi, S. (2009). New measures of well-being. *Social Indicators Research: Assessing Well-Being, 39*, 247–266. https://doi.org/10.1007/978-90-481-2354-4_12

Diener, E., Wirtz, D., Tov, W., Kim-Prieto, C., Choi, D.-W., Oishi, S., … Biswas-Diener, R. (2010). New well-being measures: Short scales to assess flourishing and positive and negative feelings. *Social Indicators Research, 97*, 143–156. https://doi.org/10.1007/s11205-009-9493-y

Du Plessis, G. A., & Guse, T. (2017). Validation of the Scale of Positive and Negative Experience in a South African student sample. *South African Journal of Psychology, 47*, 184–197. https://doi.org/10.1177/0081246316654328

Fredrickson, B. L. (1998). What good are positive emotions? *Review of General Psychology, 2*, 300–319. https://doi.org/10.1037/1089-2680.2.3.300

Gignac, G. E., & Szodorai, E. T. (2016). Effect size guidelines for individual differences researchers. *Personality and Individual Differences, 102*, 74–78. https://doi.org/10.1016/j.paid.2016.06.069

Giuntoli, L., Ceccarini, F., Sica, C., & Caudek, C. (2017). Validation of the Italian versions of the Flourishing Scale and of the Scale of Positive and Negative Experience. *SAGE Open, 7*, 1–12. https://doi.org/10.1177/2158244016682293

Harmon-Jones, C., Bastian, B., & Harmon-Jones, E. (2016). The Discrete Emotions Questionnaire: A new tool for measuring state self-reported emotions. *PLoS ONE, 11*, e0159915. https://doi.org/10.1371/journal.pone.0159915

Henry, J. D., & Crawford, J. R. (2005). The short-form version of the Depression Anxiety Stress Scales (DASS-21): Construct validity and normative data in a large non-clinical sample. *British Journal of Clinical Psychology, 44*, 227–239. https://doi.org/10.1348/014466505X29657

Howell, A. J., & Buro, K. (2015). Measuring and predicting student well-being: Further evidence in support of the Flourishing Scale and the Scale of Positive and Negative Experiences. *Social Indicators Research, 121*, 903–915. https://doi.org/10.1007/s11205-014-0663-1

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1–55. https://doi.org/10.1080/10705519909540118

Isaacowitz, D. M., & Smith, J. (2003). Positive and negative affect in very old age. *Journals of Gerontology: Psychological Sciences and Social Sciences, 58*, 143–152. https://doi.org/10.1093/geronb/58.3.P143

John, O. P., & Srivastava, S. (1999). The Big 5 trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (2nd ed., pp. 102–138). New York, NY: Guilford Press.

Jovanović, V. (2019). Measurement invariance of the Serbian version of the Satisfaction with Life Scale across age, gender, and time. *European Journal of Psychological Assessment, 4*, 555–563. https://doi.org/10.1027/1015-5759/a000410

Jovanović, V. (2015). Beyond the PANAS: Incremental validity of the Scale of Positive and Negative Experience (SPANE) in relation to well-being. *Personality and Individual Differences, 86*, 487–491. https://doi.org/10.1016/j.paid.2015.07.015

Jovanović, V., & Gavrilov-Jerković, V. (2016). The structure of adolescent affective well-being: The case of the PANAS among Serbian adolescents. *Journal of Happiness Studies, 17*, 2097–2117. https://doi.org/10.1007/s10902-015-9687-8

Jovanović, V., Gavrilov-Jerković, V., Žuljević, D., & Brdarić, D. (2014). Psihometrijska evaluacija Skale depresivnosti, anksioznosti i stresa-21 (DASS-21) na uzorku studenata u Srbiji [Psychometric evaluation of the Depression Anxiety Stress Scales-21 (DASS-21) in a Serbian student sample]. *Psihologija, 47*, 93–112. https://doi.org/10.2298/PSI1401093J

Kuppens, P., Realo, A., & Diener, E. (2008). The role of positive and negative emotions in life satisfaction judgments across nations. *Journal of Personality and Social Psychology, 95*, 66–75. https://doi.org/10.1037/0022-3514.95.1.66

Li, F., Bai, X., & Wang, Y. (2013). The scale of positive and negative experience (SPANE): Psychometric properties and normative data in a large Chinese sample. *PLoS ONE, 8*, e61137. https://doi.org/10.1371/journal.pone.0061137

Lim, N. (2016). Cultural differences in emotion: Differences in emotional arousal level between the East and the West. *Integrative Medicine Research, 5*, 105–109. https://doi.org/10.1016/j.imr.2016.03.004

Lovibond, S. H., & Lovibond, P. F. (1995). *Manual for the depression anxiety stress scales* (2nd ed.). Sydney, Australia: Psychology Foundation.

Lucas, R. E., & Gohm, C. L. (2000). Age and sex differences in subjective well-being across cultures. In E. Diener & E. M. Suh (Eds.), *Culture and subjective well-being* (pp. 291–317). Cambridge, MA: MIT Press.

Masumoto, K., Taishi, N., & Shiozaki, M. (2016). Age and gender differences in relationships among emotion regulation, mood, and mental health. *Gerontology and Geriatric Medicine, 2*, 1–8. https://doi.org/10.1177/2333721416637022

Molenaar, D., Dolan, C. V., Wicherts, J. M., & van der Maas, H. L. J. (2010). Modeling differentiation of cognitive abilities within the higher-order factor model using moderated factor analysis. *Intelligence, 38*, 611–624. https://doi.org/10.1016/j.intell.2010.09.002

Muthén, L. K., & Muthén, B. O. (2015). *Mplus user's guide* (7th ed.). Los Angeles, CA: Muthén & Muthén.

Park, N., Park, M., & Peterson, C. (2010). When is the search for meaning related to life satisfaction? *Applied Psychology: Health and Wellbeing, 2*, 1–13. https://doi.org/10.1111/j.1758-0854.2009.01024.x

Rahm, T., Heise, E., & Schuldt, M. (2017). Measuring the frequency of emotions-validation of the Scale of Positive and Negative Experience (SPANE) in Germany. *PLoS ONE, 12*, e0171288. https://doi.org/10.1371/journal.pone.0171288

Rhemtulla, M., Brosseau-Liard, P. É., & Savalei, V. (2012). When can categorical variables be treated as continuous? A comparison of robust continuous and categorical SEM estimation methods under suboptimal conditions. *Psychological Methods, 17*, 354–373. https://doi.org/10.1037/a0029315

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling. *Journal of Statistical Software, 48*, 1–36. https://doi.org/10.18637/jss.v048.i02

Rush, J., & Hofer, S. M. (2014). Differences in within- and between-person factor structure of positive and negative affect: Analysis of two intensive measurement studies using multilevel structural equation modeling. *Psychological Assessment, 26*, 462–473. https://doi.org/10.1037/a0035666

Sass, D. A. (2011). Testing measurement invariance and comparing latent factor means within a confirmatory factor analysis framework. *Journal of Psychoeducational Assessment, 29*, 347–363. https://doi.org/10.1177/0734282911406661

Saucier, G., Thalmayer, A. G., & Bel-Bahar, T. S. (2014). Human attribute concepts: Relative ubiquity across twelve mutually isolated languages. *Journal of Personality and Social Psychology, 107*, 199–216. https://doi.org/10.1037/a00364921037/a0036492

Silva, A. J., & Caetano, A. (2013). Validation of the flourishing scale and scale of positive and negative experience in Portugal. *Social Indicators Research, 110*, 469–478. https://doi.org/10.1007/s11205-011-9938-y

Soto, C. J. (2015). Is happiness good for your personality? Concurrent and prospective relations of the Big Five with subjective well-being. *Journal of Personality, 83*, 45–55. https://doi.org/10.1111/jopy.12081jopy.12081

Steel, P., Schmidt, J., & Shultz, J. (2008). Refining the relationship between personality and subjective well-being. *Psychological Bulletin, 134*, 138–161. https://doi.org/10.1037/0033-2909.134.1.138

Steger, M. F., Frazier, P., Oishi, S., & Kaler, M. (2006). The meaning in life questionnaire: Assessing the presence of and search for meaning in life. *Journal of Counseling Psychology, 53*, 80–93. https://doi.org/10.1037/0022-0167.53.1.80

Steptoe, A., Deaton, A., & Stone, A. A. (2015). Psychological wellbeing, health and ageing. *The Lancet, 385*, 640–648. https://doi.org/10.1016/S0140-6736(13)61489-0

Sumi, K. (2014). Reliability and validity of Japanese versions of the Flourishing Scale and the Scale of Positive and Negative Experience. *Social Indicators Research, 118*, 601–615. https://doi.org/10.1007/s11205-013-0432-6013-0432-6

Thompson, E. R. (2007). Development and validation of an internationally reliable short-form of the Positive and Negative Affect Schedule (PANAS). *Journal of Cross-Cultural Psychology, 38*, 227–242. https://doi.org/10.1177/0022022106297301

Tran, T. V. (2009). *Developing cross-cultural measurement*. Oxford, UK: Oxford University Press.

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. https://doi.org/10.1177/109442810031002

Vasić, A., Šarčević, D., & Trogrlić, A. (2011). Zadovoljstvo životom u Srbiji [Satisfaction with life in Serbia]. *Primenjena Psihologija, 2*, 151–177.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS Scales. *Journal of Personality and Social Psychology, 54*, 1063–1070. https://doi.org/10.1037/0022-3514.54.6.1063

History
Received July 10, 2018
Revision received March 4, 2019
Accepted March 13, 2019
Published online November 13, 2019
EJPA Section/Category Positive Psychology & Assessment

Open Data
The syntax files of all studies are available at https://osf.io/8wyka

Funding
This research has been supported by Ministarstvo prosvete, nauke i tehnološkog razvoja Republike Srbije (grant no. 179006).

ORCID
Veljko Jovanović
https://orcid.org/0000-0001-9248-2518

Veljko Jovanović
Department of Psychology
Faculty of Philosophy
University of Novi Sad
Dr Zorana Đinđića 2
21 000 Novi Sad
Serbia
vejo@uns.ac.rs
veljko.jovanovic@ff.uns.ac.rs

