Published in final edited form as:

Eur J Psychol Assess. 2003 ; 19(2): 131–141. doi:10.1027//1015-5759.19.2.131.

Factorial and construct validity of the Italian Positive and Negative Affect Schedule (PANAS)

Antonio Terracciano, **Robert R. McCrae**, and **Paul T. Costa Jr.**
National Institute on Aging, NIH, USA

Summary

This study provides evidence that an Italian version of the Positive and Negative Affect Schedule (PANAS) is a reliable and valid self-report measure. In an Italian sample (N = 600), the PANAS showed solid psychometric properties, and several American findings with the PANAS were replicated. The replicability of the PANAS factor structure was confirmed by high congruence coefficients between the American and Italian varimax solutions. Alternative models were tested with Confirmatory Factor Analysis; as in previous studies, the two-factor model achieved the best fit, but absolute fit indices varied with the estimation methods used. The independence/bipolarity issue was also explored: Positive and negative affect scales remain substantially independent after accounting for measurement error and acquiescence. Some predictions from the tripartite model of anxiety and depression were confirmed, and external correlates of the PANAS replicated those found in other languages and cultures. These analyses offer strong support for the construct validity of the Italian PANAS.

Keywords

positive and negative affect; CFA; estimation method; tripartite model of anxiety and depression; cross-cultural

Introduction

There are several reasons for the keen interest in the structure of affect and, in particular, the constructs of Positive Activation (PA) and Negative Activation (NA).[1] PA and NA are the most general dimensions that describe affective experience. They are the components of the structure of affect most often described by English language mood terms (e.g., Watson, Wiese, Vaidya, & Tellegen, 1999), and almost all descriptors that refer to the “basic” emotions (Izard, 1977) fall within the PA and NA clusters. PA and NA are the affective, emotional components of psychological or subjective well-being (SWB; Diener, Emmons, Larsen, & Griffin, 1985). Measures of PA and NA also have relevance in clinical research and practice as identifying features which distinguish anxiety from depression (Clark & Watson, 1991). Finally, PA and NA are strongly related to Extraversion and Neuroticism personality factors, respectively (Costa & McCrae, 1980;Watson & Clark, 1992), and represent core components of the two broad personality dimensions (see also Yik, Russell, Oceja, & Fernández Dols, 2000).

Address: Antonio Terracciano, Laboratory of Personality and Cognition, National Institute on Aging, NIH, 5600 Nathan Shock Drive, Baltimore, MD 21224, USA. Telephone: +01-410-5588358; Fax: +01-410-5588108. TerraccianoA@grc.nia.nih.gov.

[1]Positive Activation and Negative Activation are the labels that Watson and colleagues (Watson et al., 1999) have recently used to emphasize the activation component in their Positive and Negative Affect dimensions.

The Positive and Negative Affect Schedule (PANAS; Watson, Clark, & Tellegen, 1988) is the most frequently used instrument to assess PA and NA. The Positive Affect scale reflects the level of pleasant engagement, the extent to which a person feels enthusiastic, excited, active, and determined. The Negative Affect scale reflects a general dimension of unpleasant engagement and subjective distress that subsumes a broad range of aversive affects including fear, nervousness, guilt, and shame. The PANAS scales show excellent psychometric properties (reliability, convergent and divergent validity) and they have been translated into several languages including Estonian (Allik & Realo, 1997), German (Krohne, Egloff, Kohlmann, & Tausch, 1996), Russian (Balatsky & Diener, 1993), Spanish (Joiner, Sandin, Chorot, Lostao, & Marquina, 1997), Swedish (Hilleras, Jorm, Herlitz, & Winblad, 1998), and Turkish (Gencoz, 2000). Kercher (1992) and other researchers have used a shorter version of the PANAS scales that shows good psychometric characteristics and may be useful in situations in which brevity is important.

The present study investigates the psychometric properties of an Italian translation of the PANAS scales. The replicability of the factor structure will be tested with exploratory and confirmatory factor analyses. The effect of measurement error on the correlation between PA and NA will be explored. The relation with some external correlates will be presented and some predictions from the tripartite model of anxiety and depression will be tested (Clark & Watson, 1991).

Confirmatory Factor Analysis (CFA)

In light of the strong empirical robustness of the PANAS scales, there are surprising inconsistencies among CFA studies of the two-factor model. Researchers have adopted different strategies to achieve acceptable fit indices. Some have modified the basic model (in which two latent variables, PA and NA, are each estimated from 10 observed variables), whereas others have used different estimation methods (see Table 1).

Several different estimation procedures are available for CFA and there are statistical arguments in favor of the alternative approaches. The widely used maximum likelihood (ML) and generalized least squares (GLS) methods are based on the assumption of multinormal distributions. Therefore, deviations from normality, which occur with PANAS items (e.g., Lonigan, Hooe, David, & Kistner, 1999; see also Micceri 1989), may threaten the validity of ML and GLS significance tests (Bollen, 1989; Hu, Bentler, & Kano, 1992; but see Hu & Bentler, 1998; Olsson, Foss, Troye, & Howell, 2000). Alternative methods, less sensitive to departures from multivariate normality, have been suggested. Jöreskog and Sörbom (1989) proposed the unweighted least squares (ULS) procedure when the distributions are skewed, but caution is in order because ULS is not scale invariant, nor is it scale free (cf. Bollen, 1989, p.111 ff.). The asymptotic distribution-free (ADF; e.g., Browne, 1984) or, in LISREL terms, the weighted least squares (WLS) method, does not assume multivariate normality but requires extremely large sample sizes (see, Hu et al. 1992; Olsson et al. 2000). Finally, the Satorra-Bentler scaled test statistics, which perform scaling correction of the ML, should account for nonnormality in model fit statistics and significance testing (Hu et al., 1992). Because there is no conclusive evidence for the superiority of any single approach and because several have previously been used in PANAS analyses, we will test the models using alternative estimation methods (ML, GLS, ULS, WLS/ADF) and the Satorra-Bentler scaled correction of ML.

As reported in Table 1, some studies achieved acceptable fit after post-hoc modifications of the model. The correlations between residual errors allowed by these modifications represent the variance that the items share apart from the common latent factor. Typically this is interpreted as a common source of measurement error, but since all the 20 PANAS items are assessed by the same method, the meaning of the correlation paths among some residual errors is unclear. However, some correlated errors may be due to a high degree of overlap in item content (as for afraid and scared). In any case, such specifications must be supported by a strong substantive and/or empirical rationale (Jöreskog, 1993;MacCallum & Austin, 2000;Byrne, 2001) to avoid the risk of post hoc model fitting, and it remains necessary to cross-validate the modified model with independent sets of data. Using data recruited with state and trait time instructions, CFA will be conducted on different models representing the full and short version of the PANAS scales. Modified models proposed by other authors will be tested, without any further modifications.

Independence of PA and NA

Early factor analytic work (Nowlis & Nowlis, 1956; Borgatta, 1961; Bradburn, 1969) on self-reported affect found unipolar orthogonal factors where bipolar factors had been expected. Contrary to common sense, the correlation between positive and negative affect was surprisingly low. Since then, a large number of studies have investigated the nature of the dimensional structure of affect, enlarging comprehension of the affective phenomena. Nevertheless, controversial points are still pending.

As pointed out by Larsen and Diener (1992), the labels positive and negative affect are in part responsible for disagreements regarding the independence and bipolarity issue. The use of alternative labels, a careful selection of items, and reference to the circumplex model of affect (Russell, 1980; Watson & Tellegen, 1985) resolves part of the dispute. Indeed, many agree that within the circumplex model of affect, pleasantness (happy, content) and unpleasantness (sad, unhappy) are at the opposite ends of a bipolar dimension (Barrett & Russell, 1999; Tellegen, Watson, & Clark, 1999). The PANAS uses affect descriptors that combine pleasantness and high activation (excited, elated) and unpleasantness and high activation (nervous, upset), which instead form orthogonal dimensions (Barrett & Russell, 1999; Tellegen et al., 1999). In sum, in a two-dimensional model the independence and bipolarity concepts are not incompatible.

However, underlying issues such as the role of measurement error are still debated. Systematic and unsystematic measurement error may mask bipolarity in favor of independence. Random noise attenuates correlation coefficients between scales. Indeed, the more unreliable two scales are, the more independent they appear. Systematic measurement error, such as the acquiescent response style (yea-saying tendency), may shift negative correlations toward independence or positive correlations.

Using a modeling approach, Green, Goldman, and Salovey (1993) expressed the view that the independence of PA and NA is a statistical artifact; when systematic and unsystematic measurement errors are taken into account a largely bipolar structure emerges. Green et al. (1993) employed a multi-trait multi-method (MTMM) strategy to control both sources of error: Because different methods (e.g. adjective checklist, Likert rating scale, self- or peer-reports) are differentially affected by systematic bias (like acquiescence), and unsystematic measurement error is uncorrelated across methods, it is possible with MTMM design to partial out both sources of error. However, as pointed out by Watson et al. (1999), the MTMM approach requires that the measures are parallel forms of the same construct. “The ever-present danger is that systematic differences in content will emerge across the various formats, thereby distorting the results” (Watson et al., 1999, p. 827). Indeed, the correlation between PA and NA, as estimated by the MTMM approach, is affected by control of measurement error as well as the degree of correspondence of two or more measures used in assessing the same construct. Furthermore, later empirical evidence showed that PA and NA are relatively independent

(Watson & Clark, 1997; Watson et al., 1999) or separable (Diener, Smith, & Fujita, 1995), even after measurement error is controlled through the MTMM approach.

In this article, the relations between PA and NA will be explored with an Italian sample, providing a cross-cultural perspective on the question of independence/bipolarity. Using CFA, it is possible to estimate the correlation between the hypothesized latent factors; thus the effect of random measurement error can be partialled out. The effect of one form of systematic measurement error, acquiescence, will be addressed using an independent measure of that response style.

Method

Participants

A student sample was recruited from the Universities of Naples and Trieste. An adult sample from north and south Italy was recruited with a snowball strategy: Initial participants asked other persons (relatives, friends, partners, and acquaintances) to take part in a psychological study by completing questionnaires at home and recruiting further participants. The combined sample consisted of 600 participants (age: M = 27.9, SD = 9.78; 62.9% women). Participants were volunteers with an average to high level of education. All participants completed a written consent approved by the Johns Hopkins Bayview Medical Center's Institutional Review Board.

Measures

The Positive and Negative Affect Schedule—Participants completed the PANAS scales (Watson et al., 1988), which are composed of 10 items each. The first author completed the translation of the PANAS scales and 32 other affect items. Bilinguals blind to the content of the original English words performed back translations. The back translations were virtually identical to the original English. The translated and the English versions of the PANAS are reported in Table 3. In this Italian version the item concentrating (concentrato) was substituted for alert (allerta) because in Italian the valence of alert is ambivalent, whereas concentrating has a clear positive valence. In the PANAS-X (an expanded version of the PANAS; Watson & Clark, 1990) alert and concentrating are items of the same scale (i.e., attentiveness). This choice was supported by preliminary data. Another item, ashamed, was translated using the noun (vergogna) instead of the adjective (vergognoso), because the adjective has distinct meanings in Italian (e.g., shameful, outrageous, or shy) that might be interpreted differently by different subjects.

The short PA scale consists of the items excited, enthusiastic, concentrating, inspired and determined, whereas the short NA consists of the items distressed, upset, scared, nervous, and afraid.

The Revised NEO Personality Inventory—The NEO-PI-R (Costa & McCrae, 1992) is a 240-item questionnaire that assesses the basic dimensions of the Five-Factor Model (FFM) of personality: Neuroticism (N), Extraversion (E), Openness to Experience (O), Agreeableness (A), and Conscientiousness (C). Eight-item scales are used to measure six specific traits or facets for each of the five factors. Items are answered on a 5-point Likert scale ranging from strongly disagree to strongly agree, and scales are balanced to control for the effects of acquiescence.

A subsample of 575 partecipants filled out the Italian version of the NEO-PI-R translated by Caprara and Barbaranelli (McCrae et al., 1999), slightly modified and validated by the first author (Terracciano, 2001). For the five domain scales, the Cronbach alphas were 0.91, 0.88, 0.87, 0.86, and 0.91 for N, E, O, A, and C, respectively. These values were as high as the corresponding values for the original scales (0.92, 0.89, 0.87, 0.86, and 0.90, respectively; Costa & McCrae, 1992, Table 5). The Italian version of the NEO-PI-R closely replicates the American normative structure. Congruence coefficients comparing the Italian factors with American normative factors ranged from.96 to.98 after orthogonal Procrustes rotation (McCrae, Zonderman, Bond, Costa, & Paunonen, 1996), and all 30 facets loaded chiefly on the intended factor (Terracciano, 2001).

CES-D—About three months later, a subset of participants (n = 60) completed the affect scales for a second time. In addition, the same subset completed the Italian version (Fava, 1983) of the Center for Epidemiological Studies Depression Scale (CES-D; Radloff, 1977). The CES-D is a 20-item self-rating scale designed to assess depressive symptoms in the general population. The 20 items were selected to represent the major symptoms of depression (e.g., poor appetite, difficulty in concentrating) with emphasis on the affective components (depressed mood). Respondents reported the frequency of symptom occurrence on a four-point scale from 0 (rarely or none of the time) to 3 (most of the time or 5 to 7 days) within the last week. Four items are reverse-keyed.

Procedures

Volunteers signed the consent form, reported their address if they were interested in feedback concerning the personality profile, provided background information, and then completed questionnaires at home. They rated on a 5-point scale the extent to which they had experienced each affect term using trait and state formats. First, participants were asked to report the intensity (from slightly or not at all to extremely) of their current affect, how they were feeling “right now,” which is intended to assess state affect. Later, they were asked to report the frequency (from never to always) of their affect over an extended period of time, how they felt “in general,” which is intended to assess trait affect. Between the two measures, subjects filled out the Italian version of the NEO-PI-R.

Data analysis

Because of missing data, the following analyses regarding the PANAS scales were conducted on samples ranging from 588 to 600 participants.

Confirmatory factor analyses were conducted using the AMOS 4.01 program (Arbuckle, 1994) under SPSS (version 10.0 for window; SPSS Inc.) and EQS (Bentler, 1995). Multiple fit indices were used to evaluate the fit of hypothesized models, because different goodness-of-fit indices emphasize different aspects of model fit (Tanaka, 1993). For fit index labels and rules-of-thumb, see Table 1.

Results

Normative and reliability data

Table 2 presents means and standard deviations for PA and NA scales assessed with the momentary and general time frames. The means value are also reported separately for men and women. One-way analysis of variance (ANOVA) shows consistent sex differences. Women score significantly higher than men on the NA scale, with both state ($F(1, 591) = 8.45$, $p = .004$) and trait ($F(1, 587) = 39.65$, $p < .001$) formats. Although similar sex differences were not found in the U.S. normative sample (Watson et al. 1988), they were reported in two other studies (Lonigan et al., 1999;Mackinnon et al., 1999). These results appear consistent with commonly held beliefs (e.g., Fabes & Martin, 1991) and with sex differences reported in personality traits (e.g., Costa, Terracciano, & McCrae, 2001).

Internal consistency reliabilities are also reported in Table 2. Cronbach's coefficients α are as high as the corresponding values for the original scales (Watson et al., 1988).

Sixty subjects filled out the scales on a second occasion, about three months later. The test-retest correlations are shown in the last column of Table 2. The retest stability of the Italian scales is similar to the coefficient of the original scales. As expected, the trait format shows higher test-retest stability. However, the differences between the two formats are small, suggesting the strong influence of dispositional affect level when momentary affect is reported. Another reason that may explain the elevated stability of state affect is a possible stability in the settings and time chosen to fill out the affect measures.

Factor structure

Exploratory principal component analysis gave two clear factors corresponding to positive and negative affect. Varimax loadings for the PANAS terms for both state and trait formats are presented in Table 3. All of the descriptors have strong primary loadings on the appropriate factor, and the secondary loadings are all acceptably low.

The factor structure of the Italian PANAS scales closely replicates the original American scales. Indeed, the total congruence coefficients (McCrae et al., 1996) between Watson et al.'s original matrix (1988, p. 1067) and the Italian varimax solution were.97 and.98 for the state and trait versions, respectively. A congruence coefficient greater than.85 (Haven & ten Berge, 1977) or.90 (Mulaik, 1972; McCrae et al., 1996) is usually considered evidence of factor replication.

PANAS short—The shorter scales demonstrated adequate reliability: Cronbach's alpha was. 72 for PA and.80 for NA using the state time instruction, and.72 for PA and.83 for NA using the trait time instruction. Test-retest correlations were.77 for PA and.72 for NA using trait time instruction, and.62 for PA and.51 for NA using state time instruction. As for the full scales, two factors corresponding to PA and NA were identified through exploratory factor analysis. All descriptors loaded clearly on the expected factor.

CFA of the PANAS

An examination of the distributional properties of PA and NA scale items suggested departure from univariate normality for the state but not trait format. Several NA state items (i.e., guilty, ashamed, afraid, hostile, scared, upset, and distressed) are positively skewed and/or excessively kurtotic. Furthermore, the normalized estimate of the Mardia's (1970, 1974) coefficient of multivariate kurtosis yields significant positive kurtosis value (71.9 and 28.3 respectively for state and trait data), highly suggestive of nonnormality (Bentler, 1989; Byrne, 1994).

Oblique two-factor models—As can be seen in Table 4, using the ML method, the two-factor model is rejected according to several goodness of fit indexes, including the robust Satorra-Bentler scaled chi-square (state: S-B $\chi^2_{(169)}$ = 693.38; trait: S-B $\chi^2_{(169)}$ = 824.69, $p <.001$) and the robust CFI (state:.818; trait:.810). Using GLS and ADF/WLS estimation methods, slightly better fits were obtained according to some absolute fit indices (e.g., RMSEA, which directly measures how well an a priori model reproduces the sample data), but a worse fit according to comparative fit indices (e.g., NFI, TLI, and CFI, which assess the improvement in fit by comparing the hypothesized model against a baseline model). Finally, the model seems to achieve acceptable fit indices using the ULS estimation procedure. However, these fit indices are evaluated against cut-off values based on ML estimation, and it may possible that different cut-off values are required for GLS, ULS, and ADF estimation methods (Hu & Bentler, 1998). Across all estimation methods, only the highly recommended RMSEA (e.g., Hu & Bentler, 1998;MacCallum & Austin, 2000) is in the range of acceptable to marginal fit, with a narrow confidence interval indicating precise estimate of fit (Byrne, 2001).

Across all estimation methods, the parameters are all statistically significant. No large differences were noted on the regression weights among estimation procedures. Standardized regression weights (factor loadings) are all substantial with mean values in the range of.6 for both state and trait.

When the modified model specified by Joiner et al. (1997) was tested with the present data, 23 of the 38 covariances estimates among residual errors were not significant, indicating that the model was not parsimonious.

Alternative models—Using ML, the basic oblique two-factor model was compared to a one-factor model (which assumed that all of items pertained to the same factor), and a two-factor orthogonal model (with the correlation between PA and NA fixed to zero). The one-factor model fit the data worse than the two-factor model with both the state ($\chi^2_{(170)} = 2280.60$; CFI =.467, RMSEA =.145) and the trait ($\chi^2_{(170)} = 2274.37$; CFI =.484, RMSEA =.145) data. No substantial differences emerged between several fit indices for the oblique and orthogonal two-factor model. However, with the trait version, the oblique model fit the data significantly better than did the orthogonal model ($\Delta\chi^2_{(1)} = 7.93$; $p < .01$).

PANAS short—The pattern of differences among the estimation methods for the PANAS short was similar to that found for the full PANAS scales. A better fit emerged with the ULS and worse with ML estimator. In Table 5 we present the ML fit indices for alternative models: (a) the one-factor model; (b) the orthogonal two-factor model; (c) the basic (oblique) two-factor model; (d) the model specified by Kercher (1992), which is similar to the basic two-factor model, but the residual errors of the items scared and afraid are allowed to correlate; (e) the model specified by Mackinnon et al. (1999), which is similar to the Kercher's model but in addition the residual errors from distressed and upset are allowed to correlate.

All fit indices indicated that the one-factor model did not adequately explain the data. Although there were no substantial differences between the orthogonal and oblique two-factor models for several fit indices, the difference in χ^2 was significant (state $\Delta\chi^2_{(1)} = 5.75$; trait $\Delta\chi^2_{(1)} = 5.47$; $p < .05$). As suggested by Kercher (1992), allowing correlated errors between scared and afraid improved model fit (state $\Delta\chi^2_{(1)} = 97.23$; trait $\Delta\chi^2_{(1)} = 113.51$; $p < .001$), reaching acceptable levels according to several fit indices. In contrast, the present data do not support the additional modification suggested by Mackinnon et al. (1999).

Independence

The ML estimated correlations between the latent PA and NA factors were −.09 (95% CI = −.19 and.01) and −.14 (95% CI = −.24 and −.03) respectively for state and trait data (the short scales show results similar to the full scales in this analysis as well as the following). These values are very similar to the ones obtained between the observed PA and NA scores (state: $r = -.07$, $p > .05$; trait: $r = -.09$, $p < .05$). The trivial effect of random measurement errors is not surprising because of the high reliability of the PANAS scales. To test whether the bipolarity was compromised by methodological bias, we checked the influence of systematic error.

Acquiescence—A major known source of systematic measurement error is the acquiescent response style. The yea-saying tendency may increase the covariance of the variables, or mask inverse correlation. A partial correlation was employed to control the effect of response style, using as an indicator of acquiescence the sum of the 240 items of the NEO-PI-R. Since roughly half of the items for each factor are negatively keyed, the sum has no substantive meaning. It is, however, a content-free measure of response style, reliable and relatively stable over time (McCrae, Herbst, & Costa, 2001). High scores reflect a yea-saying tendency whereas low scores reflect the tendency to disagree regardless of item content. When controlled for this source of systematic error the correlations increase slightly for the state data ($r = -0.11$; $p < .05$), and remain virtually the same for the trait data ($r = -0.09$; $p < .05$). These analyses suggest that the independence of PA and NA is not an artifact of acquiescence response style.

External validity

Evidence of the validity of the Italian version comes from the pattern of relationship with personality and depression measures. As reported in Terracciano (2001), the relation between PA and Extraversion and NA and Neuroticism found with this Italian sample replicated the American findings (e.g., Costa & McCrae, 1980; McCrae & Costa, 1991; Watson & Clark, 1992). Consistent with expectations, PA correlated positively with Extraversion (trait: $r = .51$; $p < .01$; state: $r = .32$; $p < .01$), whereas NA was strongly related to Neuroticism (trait: $r = .69$; $p < .01$; state: $r = .38$; $p < .01$). Further, the PANAS scales appear to be good predictors of the CES-D score. In the subsample of 60 participants retested about 3 months later (Time 2), the CES-D showed strong positive correlations with NA (trait: $r = .66$; $p < .01$; state: $r = .55$; $p < .01$) and significant negative correlations with the PA scale (trait: $r = -.37$; $p < .05$; state: $r = -.27$; $p < .01$) administrated at Time 1. As expected, even higher values support concurrent validity between CES-D scores and PANAS scales administrated at Time 2. The CES-D correlated strongly with the NA scale (trait: $r = .75$; $p < .01$; state: $r = .62$; $p < .01$), and inversely with the PA scale (trait: $r = -.42$; $p < .01$; state: $r = -.56$; $p < .01$).

Tripartite model of anxiety and depression—Although anxiety and depression “share a substantial component of general distress, they can be differentiated on the basis of factors specific to each syndrome” (Clark & Watson, 1991, p. 316). The tripartite model posits that NA is the common component of general emotional distress. The other two components of the tripartite model allow differentiation of anxiety from depression: low levels of PA (anhedonia) characterize depression, whereas elevated levels of physiological hyperarousal characterize anxiety. The data from this Italian sample convey some cross-cultural support for the PA and NA portion of the Clark and Watson tripartite model (see Kiernan, Laurent, Joiner, Catanzaro, & MacLachlan, 2001). As in American studies (e.g., Lonigan et al., 1999), NA was strongly related to NEO-PI-R trait measures of both Anxiety ($r = .62$) and Depression ($r = .56$). In contrast, PA was more strongly related to Depression ($r = -.39$) than to Anxiety ($r = -.19$). This difference is statistically significant ($z = 5.52$; $p < .0001$; see Meng, Rosenthal, & Rubin, 1992). With stepwise regression analyses, NA accounts for 31% of the variance of N3: Depression ($\beta = .53$) and 38% of the variance of N1: Anxiety ($\beta = .61$). Consistent with the tripartite model, PA accounts for an additional 12% of variance of N3 ($\beta = -.34$) but only 2% of the variance of N1 ($\beta = -.14$).

Discussion

The present article provides evidence that the Italian version of the PANAS is a reliable and valid measure of self-reported affect. Indeed, solid psychometric properties and remarkable cross-cultural convergence emerged in the present study. For example, the association between PA and NA scales was weak, supporting the independence of the two factors even when the effect of measurement error was taken into account. External correlates were similar to those found in other languages and cultures (e.g., Allik & Realo, 1997). Predictions from the tripartite model of anxiety and depression were also confirmed: PA was significantly more related to depression than anxiety personality traits. This finding provides partial support for the tripartite model in another culture.

One objective of this study was the assessment of the replicability of the PANAS factor structure, and the evaluation by CFA of whether the data fit the hypothesized two-factor model. The exploratory factor analysis did produce a very clear replication of the Watson et al.

(1988) factor structure, confirmed by high values of the congruence coefficients. More problematic, but still resembling previous studies, were the results of CFA. The estimation methods employed produced different evaluations of the overall fit of the two-factor model. The basic two-factor model was rejected by the normal theory-based ML fit indices. However, unlike other studies that have found acceptable fit using WLS/ADF (i.e., Crocker, 1997) or the Satorra-Bentler scaled test statistic (i.e., Lonigan et al., 1999; see also, Sandin et al., 1999), in the present study the best fit was achieved using the ULS method. Of interest, RMSEA was in the range of acceptable to marginal fit across different estimators.

Clearer results come from the comparison of alternative models. The one-factor model never fit the data as well as the two-factor model. The modification suggested by Kercher (1992) dramatically improved the fit of the model, whereas the Joiner et al. (1997) and Mackinnon et al. (1999) modifications did not. These results recall another serious problem of studies that use CFA: the lack of objective criteria for post-hoc modification (see Vassend & Skrondal, 1997). It is a matter for future research to reduce the undesirable subjectivity in choosing the appropriate estimation method and the basis for model specification. Compared to exploratory factor analysis, CFA remains a poorly-understood technique, and provides statistical tests that may lead to a rejection of a model for unclear or trivial reasons (van de Vijver & Leung, 2001).

Many stereotypes of Italians suggest that they are characterized by affective volatility. If we assume scalar equivalence (van de Vijver & Leung, 1997; van de Vijver & Poortinga 1997; McCrae, 2001)--that raw scores can be meaningfully compared--then data from the present study offer some support for this notion. In the present Italian sample, with both the trait and the state data, PA is slightly lower and NA slightly higher than the mean reported by Watson et al. (1988). Although these effects may be due to translation inequivalencies, different response styles, or sample characteristics, they are congruent with cross-cultural study of personality traits. In fact, McCrae (2001) found that Italians score higher than Americans on the NEO-PI-R Neuroticism factor, and lower on the Extraversion factor. The validation of an Italian version of the PANAS scales is of interest for research in Italy as well as for cross-cultural research, enlarging the range of languages/cultures where the two-factor model applies, and making possible cross-cultural comparisons.

Acknowledgments

Thanks to Ivan Conte, Lisa Di Blas, Carmen Migliaccio, and Bruno Varriale for their help in recruiting subjects, and Giovanni Fava for providing the Italian version of the CES-D.

References

Allik J, Realo A. Emotional experience and its relation to the five-factor model in Estonian. Journal of Personality 1997;65:625–647.

Arbuckle, JL. AMOS (Version 4.01). Chicago: SmallWaters Corp; 1994.

Balatsky G, Diener E. Subjective well-being among Russian students. Social Indicators Research 1993;28:225–243.

Barrett LF, Russell AA. The structure of current affect: Controversies and emerging consensus. Current Directions in Psychological Science 1999;8:10–14.

Bentler, PM. EQS: Structural equations program manual. Los Angeles: BMDP Statistical Software; 1989.

Bentler, PM. EQS: Structural equations programs. Encino, CA: Multivariate Software; 1995.

Bollen, KA. Structural equations with latent variables. New York: Wiley; 1989.

Borgatta EF. Mood, personality and interaction. Journal of General Psychology 1961;64:105–137.

Bradburn, NM. The structure of psychological well-being. Chicago: Aldine; 1969.

Browne MW. Asymptotically distribution-free methods for the analysis of covariance-structures. British Journal of Mathematical & Statistical Psychology 1984;37:62–83. [PubMed: 6733054]

Browne, MW.; Cudeck, R. Alternative ways of assessing model fit. In: Bollen, KA.; Long, JS., editors. Testing structural equation models. Newbury Park, CA: Sage; 1993. p. 136-162.

Byrne, BM. Structural equation modeling with EQS and EQS/Windows. Thousand Oaks, CA: Sage; 1994.

Byrne, BM. Structural equation modeling with AMOS. Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc; 2001.

Clark LA, Watson D. Tripartite model of anxiety and depression - Psychometric evidence and taxonomic implications. Journal of Abnormal Psychology 1991;100:316–336. [PubMed: 1918611]

Costa PT Jr, McCrae RR. Influence of Extraversion and Neuroticism on subjective well- being - Happy and unhappy people. Journal of Personality and Social Psychology 1980;38:668–678. [PubMed: 7381680]

Costa, PT., Jr; McCrae, RR. Professional manual of the Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI). Odessa, FL: Psychological Assessment Resources, Inc; 1992.

Costa PT Jr, Terracciano A, McCrae RR. Gender differences in personality traits across cultures: Robust and surprising findings. Journal of Personality and Social Psychology 2001;81:322–331. [PubMed: 11519935]

Crocker PRE. A confirmatory factor analysis of the positive affect negative affect schedule (PANAS) with a youth sport sample. Journal of Sport & Exercise Psychology 1997;19:91–97.

Diener E, Emmons RA, Larsen RJ, Griffin S. The Satisfaction With Life Scale. Journal of Personality Assessment 1985;49:71–75. [PubMed: 16367493]

Diener E, Smith H, Fujita F. The personality structure of affect. Journal of Personality and Social Psychology 1995;69:130–141.

Fabes RA, Martin CL. Gender and age stereotypes of emotionality. Personality and Social Psychology Bulletin 1991;17:532–540.

Fava GA. Assessing depressive symptoms across cultures: Italian validation of the CES-D self-rating scale. Journal of Clinical Psychology 1983;39:249–251. [PubMed: 6841626]

Gencoz T. Positive and negative affect schedule: A study of validity and reliability. Turk Psikoloji Dergisi 2000;15:19–28.

Green DP, Goldman SL, Salovey P. Measurement error masks bipolarity in affect ratings. Journal of Personality and Social Psychology 1993;64:1029–1041. [PubMed: 8326466]

Haven, S.; ten Berge, JMF. Heymans Bulletin 290 EX. University of Groningen; 1977. Tucker's coefficient congruence as a measure of factorial invariance: An empirical study.

Hilleras PK, Jorm AF, Herlitz A, Winblad B. Negative and positive affect among the very old - A survey on a sample age 90 years or older. Research On Aging 1998;20:593–610.

Hu LT, Bentler PM. Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. Psychological Methods 1998;3:424–453.

Hu LT, Bentler PM, Kano Y. Can test statistics in covariance structure-analysis be trusted. Psychological Bulletin 1992;112:351–362. [PubMed: 1454899]

Izard, CE. Human emotions. New York: Plenum Press; 1977.

Joiner TE, Sandin B, Chorot P, Lostao L, Marquina G. Development and factor analytic validation of the SPANAS among women in Spain: (More) cross-cultural convergence in the structure of mood. Journal of Personality Assessment 1997;68:600–615. [PubMed: 9170299]

Jöreskog, KG. Testing structural equation models. In: Bollen, KA.; Long, JS., editors. Testing structural equation models. Newbury Park, CA: Sage; 1993. p. 294-316.

Jöreskog, KG.; Sörbom, D. LISREL 7: User's reference guide. 1. Scientific Software, Inc; 1989.

Kercher K. Assessing subjective well-being in the old-old: The PANAS as a measure of orthogonal dimensions of positive and negative affect. Research On Aging 1992;14:131–168.

Krohne HW, Egloff B, Kohlmann CW, Tausch A. Investigations with a German version of the positive and negative affect schedule (PANAS). Diagnostica 1996;42:139–156.

Kiernan G, Laurent J, Joiner TE, Catanzaro SJ, MacLachlan M. Cross-cultural examination of the tripartite model with children: Data from the Barretstown studies. Journal of Personality Assessment 2001;77:359–379. [PubMed: 11693864]

Larsen, RJ.; Diener, E. Promises and problems with the circumplex model of emotion. In: Clark, MS., editor. Review of personality and social psychology: Emotion. Vol. 13. Newbury Park, CA: Sage; 1992. p. 25-59.

Lonigan CJ, Hooe ES, David CF, Kistner JA. Positive and negative affectivity in children: Confirmatory factor analysis of a two-factor model and its relation to symptoms of anxiety and depression. Journal of Consulting and Clinical Psychology 1999;67:374–386. [PubMed: 10369058]

MacCallum RC, Austin JT. Applications of structural equation modeling in psychological research. Annual Review of Psychology 2000;51:201–226.

Mackinnon A, Jorm AF, Christensen H, Korten AE, Jacomb PA, Rodgers B. A short form of the Positive and Negative Affect Schedule: evaluation of factorial validity and invariance across demographic variables in a community sample. Personality and Individual Differences 1999;27:405–416.

Mardia KV. Measures of multivariate skewness and kurtosis with applications. Biometrika 1970;57:519–530.

Mardia KV. Applications of some measures of multivariate skewness and kurtosis in testing normality and robustness studies. Sankhva 1974;B36:115–128.

McCrae RR. Trait psychology and culture: Exploring intercultural comparisons. Journal of Personality 2001;69:819–846. [PubMed: 11767820]

McCrae RR, Costa PT Jr. Adding Liebe und Arbeit: The full Five-Factor Model and well-being. Personality and Social Psychology Bulletin 1991;17:227–232.

McCrae RR, Costa PT, de Lima MP, Simöes A, Ostendorf F, Angleitner A, Marušić I, Bratko D, Caprara GV, Barbaranelli C, Chae JH, Piedmont RL. Age differences in personality across the adult life span: Parallels in five cultures. Developmental Psychology 1999;35:466–477. [PubMed: 10082017]

McCrae, RR.; Herbst, JH.; Costa, PT, Jr. Effects of acquiescence on personality factor structures. In: Riemann, R.; Spinath, FM.; Ostendorf, F., editors. Personality and temperament: Genetics, evolution, and structure. Berlin: Pabst Science Publishers; 2001. p. 217-231.

McCrae RR, Zonderman AB, Bond MH, Costa PT Jr, Paunonen SV. Evaluating replicability of factors in the revised NEO personality inventory: Confirmatory factor analysis versus procrustes rotation. Journal of Personality and Social Psychology 1996;70:552–566.

Melvin GA, Molloy GN. Some psychometric properties of the Positive and Negative Affect Schedule among Australian youth. Psychological Reports 2000;86:1209–1212. [PubMed: 10932581]

Meng XL, Rosenthal R, Rubin DB. Comparing correlated correlation-coefficients. Psychological Bulletin 1992;111:172–175.

Micceri T. The unicorn, the normal curve, and other improbable creatures. Psychological Bulletin 1989;105:156–166.

Molloy GN, Pallant JF, Kantas A. A psychometric comparison of the positive and negative affect schedule across age and sex. Psychological Reports 2001;88:861–862. [PubMed: 11508035]

Mulaik, SA. The foundations of factor analysis. New York: McGraw-Hill; 1972.

Nowlis V, Nowlis HH. The description and analysis of mood. Annals of the New York Academy of Sciences 1956;65:345–355. [PubMed: 13373215]

Olsson UH, Foss T, Troye SV, Howell RD. The performance of ML, GLS, and WLS estimation in structural equation modeling under conditions of misspecification and nonnormality. Structural Equation Modeling 2000;7:557–595.

Radloff LS. The CES-D Scale: A self-report depression scale for research in the general population. Applied Psychological Measurement 1977;1:385–401.

Russell JA. A circumplex model of affect. Journal of Personality and Social Psychology 1980;39:1161–1178.

Sandin B, Chorot P, Lostao L, Joiner TE, Santed MA, Valiente RM. The PANAS scales of positive and negative affect: Factor analytic validation and cross-cultural convergence. Psicothema 1999;11:37–51.

Tanaka, JS. Multifaceted conceptions of fit in structural equation models. In: Bollen, KA.; Long, JS., editors. Testing structural equation models. Newbury Park, CA: Sage; 1993. p. 10-39.

Tellegen A, Watson D, Clark LA. On the dimensional and hierarchical structure of affect. Psychological Science 1999;10:297–303.

Terracciano A. The Italian version of the NEO PI-R: Conceptual and empirical support for the use of targeted rotation. Pers Individ Dif 2003;35:1859–1872. [PubMed: 19002272]

van de Vijver, FJR.; Leung, K. Methods and data analysis of comparative research. In: Berry, JW.; Poortinga, YH.; Pandey, J., editors. Handbook of cross-cultural psychology: Vol. 1: Theory and method. Boston: Allyn and Bacon; 1997. p. 257-300.

van de Vijver FJR, Leung K. Personality in cultural context: Methodological issues. Journal of Personality 2001;69:1007–1031. [PubMed: 11767816]

van de Vijver FJR, Poortinga YP. Towards an Integrated Analysis of Bias in Cross-Cultural Assessment. European Journal of Psychological Assessment 1997;13:29–37.

Vassend O, Skrondal A. Validation of the NEO Personality Inventory and the five-factor model. Can findings from exploratory and confirmatory factor analysis be reconciled? European Journal of Personality 1997;11:147–166.

Watson, D.; Clark, LA. Preliminary manual for the positive and negative affect schedule (expanded form). Southern Methodist University; Dallas: 1990. Unpublished manuscript

Watson D, Clark LA. On traits and temperament - General and specific factors of emotional experience and their relation to the 5-factor model. Journal of Personality 1992;60:441–476. [PubMed: 1635050]

Watson D, Clark LA. Measurement and mismeasurement of mood: Recurrent and emergent issues. Journal of Personality Assessment 1997;68:267–296. [PubMed: 16370781]

Watson D, Clark LA, Tellegen A. Development and validation of brief measures of positive and negative affect: The PANAS scales. Journal of Personality and Social Psychology 1988;54:1063–1070. [PubMed: 3397865]

Watson D, Tellegen A. Toward a consensual structure of mood. Psychological Bulletin 1985;98:219–235. [PubMed: 3901060]

Watson D, Wiese D, Vaidya J, Tellegen A. The two general activation systems of affect: Structural findings, evolutionary considerations, and psychobiological evidence. Journal of Personality and Social Psychology 1999;76:820–838.

Yik MSM, Russell JA, Oceja LV, Fernández Dols JM. Momentary affect in Spanish: Scales, structure, and relationship to personality. European Journal of Psychological Assessment 2000;16:160–176.

Table 1

Confirmatory Factor Analysis studies of the Positive and Negative Affect Schedule (PANAS).

Author	Estimation method	Post-hoc modifications	χ^2 (df)	Fit indices
Crocker 1997	ADF/WLS	-	551(169)	GFI=.95; AGFI=.94; RMR=.08
Joiner et al. 1997	ML	38	255(131)	GFI=.97; AGFI=.94; RMR=.04
Lonigan et al.1999	S-B	-	217(169)[a]	CFI=.94[a]; TLI=.89; RMSEA=.05
	S-B	-	208(169)[a]	CFI=.89[a]; TLI=.84; RMSEA=.06
Melvin & Molloy 2000	ML	-	902(169)	GFI=.72; AGFI=.68; RMSEA=.13
Molloy et al. 2001	ML	-	496(169)	GFI=.82; AGFI=.78; RMSEA=.09
	ML	-	932(169)	GFI=.82; AGFI=.78; RMSEA=.10
PANAS short				
Kercher 1992	ML	-	186(34)	CFI=.87; NFI=.85; NNFI=.83
	ML	1	111(33)	CFI=.93; NFI=.91; NNFI=.91
Mackinnon et al.1999	ML	-	768(34)	GFI=.94; TLI=.90; RMSEA=.09
	ML	2	295(32)	GFI=.98; TLI=.96; RMSEA=.05

Note. According to Browne & Cudeck (1993), root-mean-square error of approximation (RMSEA) values of.05 or less indicate a good fit, values in the.05 to.08 range indicate a acceptable fit, values of.08 to.10 constitute marginal fit, and values grater then.10 constitute poor fit. Values above.95 indicate good fit for the comparative fit index (CFI), and Tucker-Lewis index (TLI), known also as the Bentler-Bonett non-normed fit index (NNFI; Hu & Bentler, 1998). A cutoff value of.9 is usually used for the normed fit index (NFI), goodness-of-fit index (GFI), and adjusted-goodness-of-fit index (AGFI), but these fit indices are no longer recommended (Hu & Bentler, 1998). The smaller the root mean square residual (RMR) and the Standardized RMR (SRMR) are, the better. Hu and Bentler (1998) suggested a value of.08 for the SRMR. ML = maximum likelihood; ADF/WLS = asymptotic distribution-free/weighted least squares; S-B = Satorra-Bentler corrected maximum likelihood.

[a] Satorra-Bentler corrected fit indices.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 2

Descriptive statistics, reliabilities and retest correlations of the PANAS.

PANAS scales	Total		Women		Men		Cronbach's Alpha	Test-Retest correlation[a]
	M	SD	M	SD	M	SD		
PA STATE	27.6	7.0	27.6	7.3	27.6	6.6	.83 (.89)	.65 (.54)
NA STATE	16.0	6.2	16.5	6.5	15.0	5.5	.85 (.85)	.52 (.45)
PA TRAIT	33.0	5.9	32.9	5.8	33.2	6.1	.83 (.88)	.76 (.68)
NA TRAIT	20.9	6.5	22.1	6.4	18.7	6.1	.87 (.87)	.73 (.71)

Note. The American normative values (Watson et al., 1988) are given in parentheses. PA = Positive Affect; NA = Negative Affect.

[a] N = 60.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 3

Factor loadings and congruent coefficients of the Italian PANAS with the American loadings (Watson et al., 1988).

PANAS descriptor	State		Variable Congruence	Trait		Variable Congruence
	PA	NA		PA	NA	
Determined (determinato)	**.73**	.00	.99	**.74**	−.10	.99
Active (attivo)	**.70**	−.03	.99	**.68**	−.06	.99
Interested (interessato)	**.55**	−.15	.99	**.65**	−.03	.99
Attentive (attento)	**.63**	−.16	.97	**.65**	.02	.97
Enthusiastic (entusiasta)	**.70**	−.08	.99	**.64**	−.15	.99
Concentrating (concentrato)	**.58**	−.19	.99	**.61**	−.04	.99
Strong (forte)	**.58**	.03	.99	**.60**	−.22	.99
Inspired (ispirato)	**.66**	−.04	.95	**.59**	.04	.99
Excited (eccitato)	**.53**	.14	.90	**.58**	.14	.94
Proud (orgoglioso)	**.62**	.18	.99	**.54**	.09	.99
Afraid (impaurito)	−.08	**.73**	.98	−.06	**.75**	.99
Upset (turbato)	−.04	**.67**	.99	−.08	**.75**	.99
Nervous (nervoso)	−.03	**.78**	.99	.07	**.75**	.99
Jittery (agitato)	.06	**.80**	.99	.03	**.75**	.99
Scared (spaventato)	−.12	**.70**	.99	−.07	**.73**	.99
Distressed (angosciato)	−.05	**.69**	.99	−.11	**.73**	.99
Guilty (colpevole)	−.02	**.49**	.98	−.05	**.59**	.99
Ashamed (vergogna)	−.02	**.47**	.99	−.15	**.58**	.99
Irritable (irritabile)	.02	**.62**	.96	.09	**.56**	.92
Hostile (ostile)	−.01	**.53**	.99	.06	**.52**	.97
Factor/total congruence	.98	.97	.97	.98	.99	.98

Note. Italian translations are given in parentheses. PA = Positive Affect; NA = Negative Affect. Loadings over.40 are given in boldface.

Table 4

Goodness of fit estimations for the PANAS two-factor model.

	Estimator									RMSEA CI	
		χ^2	GFI	AGFI	SRMR	NFI	CFI	TLI	RMSEA	90% low	90% up
State	ML	968.54	.838	.799	.074	.766	.798	.773	.090	.084	.095
	GLS	599.16	.898	.873	.083	.413	.483	.418	.066	.060	.072
	ADF/WLS	624.48	.864	.831	.127	.590	.658	.616	.068	.062	.073
	ULS	811.59	.951	.940	.075	.912	--	--	--	--	--
Trait	ML	981.81	.841	.802	.075	.770	.801	.776	.091	.085	.096
	GLS	628.21	.893	.867	.081	.376	.438	.368	.068	.062	.074
	ADF/WLS	786.77	.849	.813	.107	.479	.532	.474	.079	.073	.085
	ULS	617.71	.948	.935	.075	.909	--	--	--	--	--

Note. PANAS model d.f. = 169, $p < .001$. ML = maximum likelihood; GLS = generalized least squares; ULS = unweighted least squares; ADF/WLS = asymptotically distribution-free/weighted least squares. For fit index labels and rules-of-thumb, see Table 1.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 5

Model fit indices for PANAS short

	Model	χ^2	df	GFI	AGFI	SRMR	NFI	CFI	TLI	RMSEA	RMSEA CI	
											90% low	90% up
State	One-factor	708.87	35	.767	.634	.153	.568	.577	.457	.181	.170	.193
	Two-factor orthogonal	236.67	35	.919	.873	.072	.856	.874	.837	.099	.087	.111
	Two-factor oblique	230.92	34	.921	.872	.063	.859	.877	.837	.099	.087	.112
	Kercher (1992)	133.69	33	.956	.927	.054	.918	.937	.914	.072	.060	.085
	Mackinnon et al. (1999)	131.69	32	.957	.925	.054	.920	.937	.912	.073	.060	.086
Trait	One-factor	704.29	35	.765	.630	.152	.606	.616	.506	.180	.169	.192
	Two-factor orthogonal	254.46	35	.913	.864	.075	.858	.874	.838	.103	.092	.115
	Two-factor oblique	248.99	34	.915	.862	.067	.861	.877	.837	.104	.092	.116
	Kercher (1992)	135.48	33	.954	.923	.060	.924	.941	.920	.073	.060	.086
	Mackinnon et al. (1999)	135.46	32	.954	.921	.060	.924	.941	.916	.074	.062	.087

Note. For all χ^2, $p < .001$. For fit index labels and rules-of-thumb, see Table 1.