Universal Happiness? Cross-Cultural Measurement Invariance of Scales Assessing Positive Mental Health

Article *in* Psychological Assessment · June 2016

DOI: 10.1037/pas0000353

6 authors, including:

Angela Bieda
Ruhr-Universität Bochum
10 PUBLICATIONS **4** CITATIONS
SEE PROFILE

Gerrit Hirschfeld
Hochschule Osnabrück
62 PUBLICATIONS **477** CITATIONS
SEE PROFILE

Xiao Chi Zhang
Ruhr-Universität Bochum
9 PUBLICATIONS **38** CITATIONS
SEE PROFILE

Jürgen Margraf
Ruhr-Universität Bochum
368 PUBLICATIONS **5,299** CITATIONS
SEE PROFILE

Psychological Assessment

Universal Happiness? Cross-Cultural Measurement Invariance of Scales Assessing Positive Mental Health

Angela Bieda, Gerrit Hirschfeld, Pia Schönfeld, Julia Brailovskaia, Xiao Chi Zhang, and Jürgen Margraf

Online First Publication, June 20, 2016. http://dx.doi.org/10.1037/pas0000353

Psychological Assessment

1040-3590/16/$12.00 http://dx.doi.org/10.1037/pas0000353

Universal Happiness? Cross-Cultural Measurement Invariance of Scales Assessing Positive Mental Health

Angela Bieda
Ruhr-Universität Bochum

Gerrit Hirschfeld
Osnabrück University of Applied Sciences

Pia Schönfeld, Julia Brailovskaia, Xiao Chi Zhang, and Jürgen Margraf
Ruhr-Universität Bochum

Research into positive aspects of the psyche is growing as psychologists learn more about the protective role of positive processes in the development and course of mental disorders, and about their substantial role in promoting mental health. With increasing globalization, there is strong interest in studies examining positive constructs across cultures. To obtain valid cross-cultural comparisons, measurement invariance for the scales assessing positive constructs has to be established. The current study aims to assess the cross-cultural measurement invariance of questionnaires for 6 positive constructs: Social Support (Fydrich, Sommer, Tydecks, & Brähler, 2009), Happiness (Subjective Happiness Scale; Lyubomirsky & Lepper, 1999), Life Satisfaction (Diener, Emmons, Larsen, & Griffin, 1985), Positive Mental Health Scale (Lukat, Margraf, Lutz, van der Veld, & Becker, 2016), Optimism (revised Life Orientation Test [LOT-R]; Scheier, Carver, & Bridges, 1994) and Resilience (Schumacher, Leppert, Gunzelmann, Strauss, & Brähler, 2004). Participants included German (n = 4,453), Russian (n = 3,806), and Chinese (n = 12,524) university students. Confirmatory factor analyses and measurement invariance testing demonstrated at least partial strong measurement invariance for all scales except the LOT-R and Subjective Happiness Scale. The latent mean comparisons of the constructs indicated differences between national groups. Potential methodological and cultural explanations for the intergroup differences are discussed.

Keywords: measurement invariance, cross-cultural comparison, happiness, life satisfaction, resilience

Although mental health is not solely an absence of mental disorders, most studies in clinical psychology focus exclusively on negative aspects of mental health. A comprehensive evaluation of the state of mental health, however, could not be based on assessments of clinical symptoms and risk factors alone. Positive factors should also be taken into account, such as optimism, life satisfaction, resilience, and happiness, as they strongly and independently influence the development and course of mental disorders. These positive factors can act protectively, leading to decreased rates of incidence of mental disorders (Wood & Tarrier, 2010). Despite the increased recognition of the importance of positive factors, to date they have been studied predominantly in Western nations, limiting the global generalizability of the findings. Deeper inquiry into the nature of positive protective constructs across cultures is long overdue.

The first step in examining and promoting protective factors in a cross-cultural framework is to investigate the universality and applicability of these factors across cultures (van de Vijver, 2007). Cross-cultural validity and invariance has been insufficiently examined for psychological constructs in general (Borsboom, 2006) and for positive mental health factors in particular. Participants can vary in their interpretations of certain words, their understanding of the intended meaning of a question, or their conceptualization of an entire scale due to differences in language or in cultural assumptions (Veenhoven, 1996). These divergences can be sources of bias that lead to nonequivalent constructs across groups. Cross-cultural equivalence is shown by measurement invariance across groups. Statistically, multiple group confirmatory factor analyses conducted using structural equation models are considered the gold standard for testing measurement invariance (Jöreskog, 1971). Most studies that have examined the measurement invariance of scales assessing depressive symptoms in adolescent and adult samples of different ethnicities were not able to establish full measurement invariance, indicating cultural biases and cross-

Angela Bieda, Department of Clinical Psychology and Psychotherapy, Mental Health Research & Treatment Center, Ruhr-Universität Bochum; Gerrit Hirschfeld, Department of Quantitative Methods, Osnabrück University of Applied Sciences; Pia Schönfeld, Julia Brailovskaia, Xiao Chi Zhang, and Jürgen Margraf, Department of Clinical Psychology and Psychotherapy, Mental Health Research & Treatment Center, Ruhr-Universität Bochum.

Support by the Alexander von Humboldt Professorship awarded to Jürgen Margraf by the Alexander von Humboldt Foundation is gratefully acknowledged. We want to highly thank Holger Schillack, Helen Vollrath, Simon Blackwell, Jan Christopher Cwik, Felicity Jensz, and Julia Velten who helped in proofreading the article.

Correspondence concerning this article should be addressed to Angela Bieda, Department of Clinical Psychology and Psychotherapy, Mental Health Research & Treatment Center, Ruhr-Universität Bochum, Massenbergstr. 9-13, 44787 Bochum. E-mail: angela.bieda@rub.de

cultural differences (Baas et al., 2011; Crockett, Randall, Shen, Russell, & Driscoll, 2005; Guerra, Ferri, Llibre, Prina, & Prince, 2015; Losada et al., 2012; Nuevo et al., 2009; Wu et al., 2012). These analyses concluded that observed means should be compared across groups with caution, as they do not always reflect true latent factor mean differences.

To address the lack of focus on positive factors and their invariant measurement across cultures, this study examines social support, happiness, life satisfaction, positive mental health, optimism, and resilience, which are positive constructs considered to be among the most relevant to mental health and illness. We chose short unidimensional scales known to be valid and reliable covering a broad range of well-being dimensions. The selection of scales was roughly guided by Keyes' classification of emotional, psychological, and social well-being (Keyes, 2002). However, we did not rely on a specific theory of well-being because there is no general agreement about how best to conceptualize this construct (Wissing & Temane, 2008).

The role of perceived social support as a protective factor for mental disorders is well established (Cohen, 2004). Aside from reducing the incidence of mental illness, perceived social support also reduces general stress (Lindorff, 2000) and has positive effects on somatic health (Uchino, Cacioppo, & Kieholt-Glaser, 1996). In her cross-cultural research, Nover (2012) tested a long version of the Questionnaire-Social Support (F-SozU-22; Fydrich, Sommer, Menzel, & Höll, 1987) among pupils in Germany, Luxembourg, and Spain, finding partial weak measurement invariance for the three cultural groups. Further analyses revealed that the F-SozU-22 was invariant among the German and Luxembourgian pupils, and therefore mean comparison was possible. However, strong measurement invariance between the German and Spanish samples could not be established, and thus their means could not be compared. A short form of the Questionnaire-Social Support (F-SozU K-14; Fydrich, Sommer, Tydecks & Brähler, 2009) was adjusted to be more valid and reliable than the F-SozU-22. Certain items, including the items found to be variant were then rephrased (e.g., Item 5 and Item 14; Nover, 2012). Measurement invariance of the F-SozU K-14 has not been tested.

Traditionally, *happiness* is defined by the experience of more frequent positive affective states than negative ones (Bradburn, 1969) and by one's overall quality of life (Diener, Suh, Lucas, & Smith, 1999). However, Lyubomirsky and Lepper (1999) argued that a subjective rating of whether one is a happy or an unhappy person reflects a broader and more molar perspective and developed the Subjective Happiness Scale (SHS) to measure happiness more directly. This subjective happiness rating correlates with self-esteem, optimism, purpose in life, and global life satisfaction (Lyubomirsky, Tkach, & DiMatteo, 2006). Happiness as measured by the SHS has been examined in prior studies cross-culturally via direct group comparisons (Swami et al., 2009). However, mean comparisons can be biased when the assumption of measurement invariance is not met, resulting in potentially misleading conclusions. To date, the measurement invariance of the SHS remains untested.

Life satisfaction is defined subjectively as an overall "conscious cognitive judgment of one's life in which the criteria for judgment are up to the person" (Pavot & Diener, 2008). Life satisfaction is negatively associated with psychological distress (Marum et al., 2014) and has been shown to be an indicator and predictor of functioning and presence of clinical symptoms and comorbidity in college students (Renshaw & Cohen, 2014). Life satisfaction is mostly measured by the unidimensional Satisfaction with Life Scale (SWLS; Diener et al., 1985). Tucker, Ozer, Lyubomirsky, and Boehm (2006) evaluated the measurement invariance of the SWLS in small student and community samples in the United States and Russia, with results indicating strong measurement invariance in student samples but only weak measurement invariance in community samples. The SWLS scores were not comparable across the U.S. and Russian community-based groups. Oishi (2006) conducted a study with American and Chinese students, applying an item response theory analysis, which revealed that Items 4 and 5 assessing "ones satisfaction with past accomplishments" showed an item bias across the two cultural groups. In another study, Zanon, Bardagi, Layous, and Hutz (2014) examined measurement invariance between students from the United States and Brazil using multiple group confirmatory factor analysis and also found that Items 4 and 5 of the SWLS were noninvariant. Our study aimed to test the SWLS's measurement invariance in larger samples from three countries.

The concept of positive mental health combines mainly emotional, but also psychological and social aspects of well-being into a single general construct. People with good mental health tend to have stable relationships, view their lives as having purpose and direction, experience more positive affect, and are more likely to be self-accepting (Keyes, 2002). The Positive Mental Health Scale assesses particularly the emotional components of positive mental health (Lukat et al., 2016). Partial scalar measurement invariance has been established in student and patient samples in Germany and research on this scale is ongoing.

Optimism is a positive orientation toward the future, according to the theory outlined by Scheier and Carver (1985). It is considered a personality trait and a psychological resource, supported by research indicating associations with improved well-being and physical health (Carver, Scheier, & Segerstrom, 2010; Rasmussen, Scheier, & Greenhouse, 2009). Optimism as measured by the Life Orientation Test (Scheier et al., 1994) is a universally applicable construct across the 142 countries studied by Gallagher, Lopez, and Pressman (2013), with health advantages not limited to Western nations. However, cross-cultural measurement invariance of the revised Life Orientation Test (LOT-R) still awaits assessment.

There are several definitions of resilience, with researchers not yet in agreement as to whether resilience is an innate personality trait or a dynamic quality that emerges only in life when a stressor is present (Fletcher & Sarkar, 2013). From a clinical viewpoint, resilience is personality characteristic that buffers the negative effects of stress on mental and somatic health and promotes positive psychosocial adaption (Wagnild & Young, 1993). It is a multidimensional construct that involves different characteristics, including among others equanimity, perseverance, and self-reliance. The effectiveness of resilience interventions and training programs indicates that resilience can be promoted and can increase, and that it reduces depressive symptoms (Brunwasser, Gillham, & Kim, 2009). Resilience is typically assessed using the short form (RS-14) or long version (RS-25) of the Resilience Scale (Wagnild & Young, 1993). An 11-item version of this scale has been validated for use in Germany (Schumacher et al., 2005). An examination of measurement invariance of the RS-14 indicated partial strong invariance across student samples from the U.S.,

Taiwan and China. Latent mean comparisons were thus based on four remaining invariant items, and revealed the highest latent factor means in the U.S. student sample (Yang, Li, & Xia, 2012).

Given the explanatory and predictive power of positive constructs regarding mental health and illness outcomes, it is unfortunate that information on cross-cultural measurement invariance is available for only a few positive constructs. For example, standard measures of depression provide a great deal of cross-cultural invariance data (Baas et al., 2011; Crockett et al., 2005; Dere et al., 2015; Losada et al., 2012; Kendler et al., 2015; Skriner & Chu, 2014; Whisman, Judd, Whiteford, & Gelhorn, 2013; Zhang et al., 2011). Cross-cultural measurement invariance data for positive constructs would provide a basis for integrating positive and negative aspects of mental health in cross-cultural clinical psychology research. The present study aimed to examine whether six primary positive constructs are measurement invariant across student samples from Germany, Russia, and China. Configural invariance implies that the construct exists across cultures and that the form of the scale is equal across cultures. Weak measurement invariance refers to equal loadings and implies that structural relationships between latent variables can be meaningfully compared across cultures. Strong measurement invariance refers to equal intercepts across cultures. It permits the comparison of relationships between variables, but also comparisons of means across cultures (Hirschfeld & von Brachel, 2014; Meredith, 1993). We expected to find at least configural invariance. We further expected, based on the findings of the SWLS and RS-14, that it may be more difficult to establish more stringent weak and strong cultural measurement invariance standards, as all scales were developed in Western countries, with differences in items across the cultures being likely. Given that at least partial strong measurement invariance is met by a construct, latent means will be compared.

Method

Participants and Procedure

Participants were recruited within the longitudinal project BOOM (Bochum Optimism and Mental Health Studies), which aims to identify protective factors related to positive mental health across several countries. Participants gave their consent after being assured of anonymity and of the voluntary nature of the study. The Ethics Committee of the Faculty of Psychology of the Ruhr-Universität Bochum approved the study. All participants were students in Germany, Russia, or China. The German sample consisted of 4,240 students at the Ruhr-Universität Bochum. German participants were recruited by an e-mailed invitation with a link leading to an online questionnaire. The link was sent to all students enrolled at Ruhr-Universität Bochum. Participants in Germany were offered an incentive to participate, namely the opportunity to take part in a draw for a gift coupon or a tablet computer. The Russian sample consisted of 3,745 students from Lomonossov University Moscow, University of Voronesh, and University of Orenburg. Participants were recruited via an invitation letter. The response rate was 95.3%. Data were gathered by online and paper-and-pencil questionnaires administered in a group testing session. Participants received no financial compensation. The Chinese sample consisted of 12,524 university students from Capital Normal University Beijing, Hebei United University, and Nanjing University. Participants, mainly first-year students, were recruited during their first month of study via an invitation e-mail. The response rate was 94.5%. Data was gathered by an online or a paper-pencil questionnaire administered in a group testing session. Participants received no financial compensation. Participants' characteristics are displayed by nation in Table 1. Gender ratio, mean age, and mean number of semesters (1 semester = 0.5 academic year) studied at university differed significantly among the samples. We examined the influence of the sociodemographic factors on the items used in the analyses. The median correlation between age and the individual items was −.05. Similarly, median correlation between gender and individual items was .03. The median correlation between semester and items was .03. Thus, the sociodemographic variables had rather little effect.

Measures

Validated German versions exist for all questionnaires used in the analysis. Russian and Chinese versions of the measures were developed by using the translation-back-translation method recommended by Brislin (1970). In cases of discrepancies, this procedure was repeated by the researchers until an agreement was reached.

Questionnaire-Social Support. Social support (for an overview of all measures and internal consistencies, see Table 2) was assessed using the 14-item Questionnaire-Social Support measuring perceived and/or anticipated social support (F-SozU K-14; Fydrich et al., 2009). Participants indicated agreement with statements such as "I experience a lot of understanding and security from others" on a 5-point Likert scale ranging from 1 (*not true*) to 5 (*true*). In a German sample this unidimensional measure showed excellent Cronbach's α and good convergent and discriminant validity (Fydrich et al., 2009).

SHS. Global subjective happiness was assessed using the four-item SHS (Lyubomirsky & Lepper, 1999). Participants responded on a 7-point Likert scale whose wording of anchor points depended on the question. Responses are averaged for an overall score where high scores indicate high subjective happiness. In

Table 1
Demographic Characteristics by Country

Variable	Germany (n = 4,453)	Russia (n = 3,806)	China (n = 12,524)
Age (Mean ± *SD*; Range)	26.54 ± 4.0; 18–60	20.3 ± 2.4; 14–42	19.73 ± 1.86; 14–42
Sex [male] (%)	2,445 (54.9%)	2,479 (65.1%)	7,773 (62.1%)
Semester (M ± *SD*; range)	8.37 ± 3.45; 1–14	2.66 ± 1.45; 1–6	1.75 ± 1.2; 1–8

Note. Differences between the samples were significant, $p < .001$.

Table 2
Means, Standard Deviations, Skewness, Kurtosis, and Internal Consistency Across Scales and Countries

	German sample					Russian sample					Chinese sample				
Scale	*M*	*SD*	Skew	Kurt	α	*M*	*SD*	Skew	Kurt	α	*M*	*SD*	skew	kurt	α
F-SozU K-14	59.49	10.96	−1.49	2.22	.95	58.00	11.3	−1.5	2.2	.94	56.7	12.09	−1.28	1.46	.95
SHS	18.88	5.39	−.55	−.33	.87	20.41	4.4	−.4	−.04	.73	21.75	4.28	−.67	.45	.74
SWLS	25.08	6.55	−.79	−.1	.89	24.38	5.77	−.53	−.03	.83	23.96	6.46	−.54	−.22	.88
PMH-scale	18.09	5.97	−.54	−.3	.93	19.03	5.11	−.7	.26	.87	21.14	5.06	−.94	.93	.9
LOT-R	8.47	3.59	.01	−.12	.83	8.22	5.11	.5	−.17	.77	8.47	3.59	.01	−.12	.42
RS-10	51.03	12.85	−1.08	.06	.93	54.16	8.41	−1.31	3.2	.82	53.34	8.03	−.67	1.57	.8
RS-11	55.8	13.87	−1.07	.62	.93	59.27	8.75	−1.41	4.01	.8	58.53	8.41	−.68	1.93	.79

Note. F-SozU K-14 = Questionnaire Social Support (possible min and max values = 14–70); SHS = Subjective Happiness Scale (4–28); SWLS = Satisfaction with life Scale (5–35); PMH-scale = Positive Mental Health Scale (0–27); LOT-R = Life Orientation Test (0–24); RS-10 = Resilience Scale (10 items, removing item 10 from RS-11; this scale was used for our study; 10–70); RS-11 = Resilience Scale (11 items; 11–70); Skew = Skewness; Kurt = Kurtosis.

several countries internal consistency has been found to be good, as well as convergent and discriminant validity (Extremera & Fernandez-Berrocal, 2014; Iani, Lauriola, Layous, & Sirigatti, 2014; Swami et al., 2009).

SWLS. Satisfaction with one's life as a whole was measured using the five-item SWLS (Diener et al., 1985; Glaesmer, Grande, Braehler, & Roth, 2011). Participants indicated agreement with statements such as "In most ways my life is close to my ideal" on a scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*). The SWLS has good psychometric properties, shows convergent and discriminant validity in various samples (Pavot & Diener, 2008) and strong invariance regarding gender. Research in populations including samples from the United States, China, Brazil, and Russia did not demonstrate cross-cultural invariance (Oishi, 2006; Tucker et al., 2006; Zanon et al., 2014).

Positive Mental Health Scale. The nine-item Positive Mental Health Scale (PMH-scale; Lukat et al., 2016) assesses mainly emotional aspects of well-being without explicitly referring to well-being theories. It was developed to assess a single holistic concept of positive emotionality related to positive mental health. The PMH-scale's items derived from the Trierer Personality Inventory (Becker, 1989), the Freiburg Personality Inventory (Fahrenberg, Selg, & Hampel, 1989), Mental Health Scale (Tönnies, Plöhn, & Krippendorf, 1996), the Berner Well-Being Inventory (Grob et al., 1991), and the new Item 2 formulated by the authors. Participants respond to statements such as "I am in good physical and emotional condition" on a scale ranging from 0 (*I disagree*) to 3 (*I agree*). It has strong psychometric properties. In various groups (student and different patient samples), PMH-scale showed a unidimensional structure. The 1-week test–retest reliability is .81, its internal consistency ranged from .84 (patient sample) to .93 (student sample). In a major validation study in student samples, the PMH-scale showed convergent validity with the SHS (r = .81) and the F-SozU-22 (r = .52). Discriminant validity was measured by the Depression, Anxiety and Stress Scales (Lovibond & Lovibond, 1995) in student and patient samples. The PMH-scale correlated significantly negative with its subscales depression (r = −.74), anxiety (r = −.51), and stress (r = −.56). Research in various German samples using maximum likelihood estimation indicated invariance across various groups and time (Lukat et al., 2016).

LOT-R. Dispositional optimism was assessed with the LOT-R (Glaesmer et al., 2008; Scheier et al., 1994). The LOT-R consists of 10 items, of which three assess optimism, three assess pessimism, and the remaining four serve as filler items. Responses are indicated on a scale ranging from 0 (*I strongly agree*) to 4 (*I strongly disagree).* The LOT-R showed convergent validity with measures of coping and well-being, and discriminant validity from measures of symptom reporting and negative affect (Andersson, 1996; Monzani, Steca, & Greco, 2014). Consensus in the research indicates that the most reliable way to score the questionnaire is to average the responses to the optimism and reverse-scored pessimism items for a total score. Different models of this questionnaire have been proposed with several later studies finding that the LOT-R is best represented by a bidimensional model with one *optimism* factor loaded by all items and a method factor for the positively worded items. This model showed superiority to models in which one factor formed by the positively worded items reflecting optimism and one factor formed by the inverse items reflecting pessimism are correlated or a unidimensional model where all items loaded on one factor (Alessandri et al., 2010; Herzberg, Glaesmer, & Hoyer, 2006; Monzani et al., 2014; Rauch, Schweizer, & Moosbrugger, 2007).

Resilience Scale. Psychosocial stress resilience was assessed with an 11-item short version of the Wagnild and Young Resilience Scale (RS-11; Schumacher et al., 2005). Participants responded to items such as "I usually manage one way or another" on a scale ranging from 1 (I disagree) to 7 (I agree). The RS-11 demonstrated a good reliability and convergent validity in a German sample (Schumacher et al., 2005).

Data Analysis

Data were screened for response sets and missing values. Only cases with complete data were included in the analysis. Data were also screened for skewness and kurtosis (Kline, 2010). Demographic variables were examined using standard descriptive statistics within each sample, including means, standard deviations, and internal consistencies. The main analysis was conducted in three steps. First, single-group confirmatory factor analyses (CFA) examined each instrument's proposed factor structure. For the scales for which various models had been proposed (e.g., the LOT-R), fit

indices of the models were compared. In case of model misspecifications, the cause of specification error was examined via modification indices. Model modifications were used with caution, and applied only if a reasonable theoretical explanation existed. At least three models, one for each country, were tested for each questionnaire. To evaluate the goodness of fit for the models, we used the fit indices and cut-off values recommended by Hu and Bentler (1999). Because the χ^2 is sensitive to sample size, the root mean square of approximation (RMSEA; Steiger, 1990) was also used for model evaluation. RMSEA values smaller than .08 indicated a reasonable fit and values smaller than .05 a good fit (MacCallum, Browne, & Sugawara, 1996). A 90% confidence interval was also reported. For the comparative fit index (CFI; Bentler, 1990), values greater than .9 indicated a good fit. For the standardized root-mean-square residual (SRMR), values smaller than .09 indicated a good fit. Factor loadings were also examined, with a minimum for factor loadings set at .40 (Ford, MacCallum, & Tait, 1986). Maximum likelihood (ML) estimation was used for model testing for all but the PMH-scale, for which we used the Weighted least squares mean and variance adjusted (WLSMV; Muthén & Asparouhov, 2002) estimator. WLSMV is recommended if fewer than five response categories are given (Beauducel & Herzberg, 2006).

Second, measurement invariance (MI) testing included a series of model comparisons. At each comparison step, equality constraints were added consecutively to the models (Steenkamp & Baumgartner, 1998, see also Meredith, 1993). Each model served as a basis for comparison to the preceding model. Measurement testing took place in the framework of multiple group CFA. The data analysis followed the procedure recommended by Vandenberg and Lance (2000). In the baseline model (configural invariance), no equality constraints were made. This enabled an evaluation of whether factor structures were the same across samples. For testing weak invariance, factor loadings were constrained to be equal across groups. If weak invariance was established, intercepts were also constrained to be equal across the groups (strong invariance). Imposing equality constraints on models will always lead to decreases in fit. To determine if the decrease in fit is substantial, initial studies used chi-square difference tests (Byrne, Shavelson, & Muthén, 1989), but chi-square differences were also sensitive to sample size (Oishi, 2007). Therefore, Cheung and Rensvold (2002) recommended a ΔCFI smaller or equal than .01 to indicate measurement invariance. In addition to the chi-square difference test, ΔCFI was examined because of the sensitivity of chi-square statistics. For the PMH-scale, so far ΔCFI seems to be the only option for models estimated with WLSMV and large sample sizes (Hirschfeld et al., 2014; Sass, Schmitt, & Marsh, 2014).

If, at a particular step, full measurement invariance could not be established, partial invariance was examined (Byrne et al., 1989; Steenkamp & Baumgartner, 1998). To test partial measurement invariance, first misspecified items were identified by means of modification indices and were then allowed to differ between groups. The order in which the parameters were released depended on the highest value of the modification indices. At least two loadings or intercepts had to be equal across groups to establish partial measurement invariance (Byrne et al., 1989).

Third, if strong invariance or partial strong invariance was established, latent means were compared. To examine differences among latent means, one latent mean of a group was fixed to zero, and the two other latent means were freely estimated and tested for significant differences. Effect sizes were reported. Analyses were conducted with the software program R and its package LAVAAN (Rosseel, 2012).

Results

Descriptive Statistics

Means, standard deviations, skewness, and kurtosis of the sum scores and internal consistencies for each scale are displayed in Table 2. None of the sum scores exhibited relevant skewness (>3) or kurtosis (>8) that exceed the cutoff values, indicating univariate normality. The internal consistency for all scales was at least good ($\alpha > .7$), except for the Chinese version of the LOT-R ($\alpha = .42$). Correlation matrices were examined to determine whether items on each scale correlate with each other significantly in the expected way. Item 10 of the RS-11 showed an unexpected mixed pattern of negative and noncorrelations with other items of the scale in the Russian and Chinese samples. Therefore, this item was excluded from further analyses, and a 10-item Resilience Scale (RS-10) was used for further analysis. Item 3 and Item 7 in the Chinese version of the LOT-R, assessing pessimism, did not correlate in the expected way with the items assessing optimism. Therefore, confirmatory factor and measurement invariance analyses for the LOT-R could only be conducted in the German and Russian samples.

Single Group Confirmatory Factor Analyses

Questionnaire-Social Support. The unidimensional models and additional model specifications for the F-SozU K-14 are displayed in Table 3. The unidimensional model of the F-SozU K-14 resulted in a reasonable CFI and SRMR; however, the RMSEA indicated a poor fit in all three samples. Modification indices suggested that there was a correlated error term between Item 10 ("If I am down, I know to whom I can go without hesitation") and Item 11 ("There are people who share joy and sorrow with me") in all three groups. Both items indicate that participants have someone to go to in case of negative or extreme feelings. After allowing the error terms to correlate, model fit improved. However, the RMSEA in the German sample still indicated a poor fit. There were more correlated error terms in the German sample, but not in the Russian and Chinese sample. All factor loadings for all three groups were at least .5.

SHS. The unidimensional models for the SHS and additional model specifications are displayed in Table 4. The proposed unidimensional model of the SHS resulted in an excellent model fit in the Russian and Chinese samples, with the RMSEA in the German sample indicating misspecification. Modification indices revealed correlated error terms between Items 1 and 2 as well as Items 3 and 4. To facilitate comparability across groups, no correlated error terms were included in the model that was tested for measurement invariance. All factor loadings for all three groups were above .4, except for Item 4 of the Russian version ($\lambda = .35$).

SWLS. The unidimensional models and additional model specifications for the SWLS are displayed in Table 5. The fit of the

Table 3
Summary of Fit Indices From Comparative Factor Analysis (CFA) and Invariance Analyses Across Samples for the F-SozU K-14

Model—F-SozU K-14	χ^2 (*df*)	RMSEA [90% CI]	CFI	SRMR	ΔCFI	$\Delta\chi^2$
Single group CFA-original one factor model						
German sample	4387.186 (77)	.112 [.109, .115]	.909	.046		
Russian sample	2196.478 (77)	.085 [.082, .088]	.934	.037		
Chinese sample	8632.551 (77)	.094 [.093, .096]	.939	.036		
Single group CFA ($\Theta_{10,\ 11}$ free)						
German sample	4017.999 (76)	.108 [.105, .111]	.917	.045		369.19
Russian sample	2058.379 (76)	.083 [.080, .086]	.939	.036		138.1
Chinese sample	7573.635 (76)	.089 [.087, .090]	.942	.035		1058.9
Multiple group CFA models						
Configural	13650.013 (228)	.092 [.091, .094]	.936	.035		
Weak	14398.235 (256)	.089 [.088, .091]	.932	.069	.004	
Strong	19347.759 (282)	.099 [.098, .100]	.909	.076	.023	
Partial strong						
$\tau_{3\ \text{free}}$	18958.225 (280)	.096 [.095-.097]	.915	.074	.017	
$\tau_{3,}\ \tau_{8\ \text{free}}$	16776.968 (278)	.093 [.091-.094]	.921	.073	.011	
$\tau_{3,}\ \tau_{8,}\ \tau_{11\ \text{free}}$	16369.580 (276)	.092 [.091-.093]	.923	.072	.009	

Note. F-SozU K-14 = Questionnaire Social Support; τ_3 = "I experience a lot of understanding and security from others"; τ_8 = "I have friends and family who will simply just hug me"; τ_{11} = "There are people who share joy and sorrow with me"; RMSEA = root mean square error of analysis; CI = confidence interval; CFI = comparative fit index; SRMR = standardized root mean square residual. all χ^2- tests and $\Delta\chi^2$ were significant, $p < .001$.

unidimensional model of the SWLS was acceptable in all groups. All factor loadings were above .5.

PMH-scale. The unidimensional models and additional model specifications for the PMH-scale are displayed in Table 5. For the PMH-scale, the global model fit was acceptable in all three groups. Again, the RMSEA for the German and Chinese sample indicated several misspecifications. In all three samples, modification indices indicated that error terms of Items 1 ("I am often carefree and in good spirits") and 2 ("I enjoy my life") were correlated. These items have the common theme of enjoyment, to live one's life with ease. Allowing for these correlations improved the model fit. The RMSEA value was now also in an acceptable range. Factor loadings exceeded the recommended minimum of .4.

LOT-R. The unidimensional models, additional model specifications, and alternative models for the LOT-R are displayed in Table 6. For all groups, three different models were compared: a unidimensional model, a bidimensional model with correlated factors reflecting optimism and pessimism, and a bidimensional model with an optimism factor and an orthogonal response factor. The unidimensional model had a poor fit in both samples, while the correlated factor and the orthogonal factor models showed an acceptable fit. The latter model performed slightly better than the model with two factors

Table 4
Summary of Fit Indices From Comparative Factor Analysis (CFA) and Invariance Analyses Across Samples for the Subjective Happiness Scale (SHS)

Model—SHS	χ^2 (*df*)	RMSEA [90% CI]	CFI	SRMR	ΔCFI
Single group CFA-original one factor model					
German sample	68.064 (2)	.086 [.069, .104]	.993	.012	
Russian sample	10.065 (2)	.033 [.015, .054]	.998	.008	
Chinese sample	50.974 (2)	.044 [.034, .055]	.997	.011	
Multiple group CFA models					
Configural	129.102 (6)	.054 [.046, .063]	.996	.011	
Weak	780.489 (14)	.089 [.084, .094]	.973	.099	.023
Partial weak					
λ_3 free	521.765 (12)	.078 [.073, .84]	.982	.074	.014
$\lambda_{3,}\ \lambda_4$ free	281.419 (10)	.063 [.056, .069]	.991	.048	.005
Strong	892.147 (16)	.089 [.084, .094]	.969	.056	.022
Partial strong					
$\tau_{3\ \text{free}}$	870.384 (14)	.094 [.089, .099]	.970	.055	.021
$\tau_{3,}\ \tau_4$ free	744.132 (12)	.094 [.088, .100]	.974	.051	.017

Note. SHS = Subjective Happiness Questionnaire; RMSEA = root mean square error of analysis; CI = confidence interval; CFI = comparative fit index; SRMR = standardized root mean square residual; λ_3 = "Some people are generally very happy. They enjoy life regardless of what is going on, getting the most out of everything. To what extent does this characterization describe you?"; λ_4 = "Some people are generally not very happy. Although they are not depressed, they never seem as happy as they might be. To what extend does this characterization describe you?" All χ^2 tests and $\Delta\chi^2$ were significant, $p < .001$.

Table 5
Summary of Fit Indices From Comparative Factor Analysis (CFA) and Invariance Analyses Across Samples for the SWLS and for the PMH-Scale

Model	χ^2 (*df*)	RMSEA [90% CI]	CFI	SRMR	WRMR	ΔCFI	Δ χ^2
Model—SWLS							
Single group CFA-original one factor model							
German sample	49.742 (5)	.045 [.034, .057]	.996	.012			
Russian sample	79.106(5)	.062 [.051, .075]	.989	.021			
Chinese sample	466.015 (5)	.086 [.079, .092]	.987	.026			
Multiple group CFA models							
Configural	594.864(15)	.075 [.070, .080]	.989	.019			
Weak	1380.089 (25)	.088 [.085, .092]	.974	.051		.015	
Partial weak							
λ_2 free	811.507 (23)	.070 [.066, .75]	.985	.051		.004	
Strong	2322.243 (29)	.107[.103, .111]	.957	.066		.028	
Partial strong							
τ_1 free	1493.641 (27)	.089 [.085-.092]	.972	.057		.013	
τ_1, τ_3 free	830.171 (25)	.068 [.064-.072]	.985	.051		.000	
Model—PMH-scale							
Single group CFA-original one factor model							
German sample	991.652 (27)	.090 [.085, .094]	.990		2.645		
Russian sample	432.687 (27)	.063 [.058, .068]	.985		1.857		
Chinese sample	3486.410 (27)	.101 [.098, .104]	.974		4.874		
Single group CFA ($\Theta_{1,2}$ free)							
German sample	706.196 (26)	.077 [.072, .082]	.993		2.203		250.56
Russian sample	319.929 (26)	.055 [.049, .060]	.989		1.589		103.82
Chinese sample	2536.871 (26)	.088 [.085, .091]	.981		4.077		863.4
Multiple group CFA models							
Configural	3364.944 (78)	.078 [.076, .080]	.988		4.899		
Weak	3739.852 (94)	.075 [.073, .077]	.986		6.580	.002	779.64
Strong	6519.005 (128)	.085 [.083, .087]	.976		7.816	.010	765.23

Note. SWLS = Satisfaction With Life Scale; PMH-scale = Positive Mental Health Scale; RMSEA = root mean square error of analysis; CI = confidence interval; CFI = comparative fit index; SRMR = standardized root mean square residual; λ_2 = "The conditions of my life are excellent"; τ_1 = "In most ways my life is close to my ideal"; τ_3 = "I am satisfied with my life." All χ^2 tests and $\Delta\chi^2$ were significant, $p < .001$.

representing optimism and pessimism. Except for Item 9 ($\lambda = .35$) in the Russian sample, all factor loadings were above .4.

Resilience Scale. The unidimensional models and additional model specifications for the RS-10 are displayed in Table 7. The global fit for the unidimensional model of the RS-10 in the German and Chinese sample was poor. Modification indices indicated several misspecifications, especially in the German sample. All three groups showed correlated error terms between the item

Table 6
Summary of Fit Indices Comparative Factor Analysis (CFA) and Invariance Analyses Across German and Russian Students for the LOT-R

Model—LOT-R	χ^2 (*df*)	RMSEA [90% CI]	CFI	SRMR	ΔCFI
Single group CFA-one factor model					
German sample	1111.349 (9)	.166 [.158, .174]	.883	.058	
Russian sample	1142.692 (9)	.182 [.173, .191]	.808	.080	
Single group CFA-two correlated factors model					
German sample	215.031 (8)	.076 [.068, .085]	.978	.033	
Russian sample	143.251 (8)	.067 [.057, .076]	.977	.034	
Single group CFA-two orthogonal factors (trait and method) model					
German sample	123.634 (5)	.073 [.062, .084]	.987	.021	
Russian sample	83.904 (5)	.064 [.053, .077]	.986	.022	
Multiple group CFA models					
Configural	207.538 (10)	.069 [.061, .078]	.987	.021	
Weak	274.570 (17)	.061 [.054, .067]	.983	.029	.004
Strong	593.647 (21)	.081 [.076, .87]	.963	.041	.024
Partial strong					
τ_3 free	359.530 (20)	.064 [.058, .070]	.978	.031	.009

Note. LOT-R = Life Orientation Test—Revised; RMSEA = root mean square error of analysis; CI = confidence interval; CFI = comparative fit index; SRMR = standardized root mean square residual; τ_3 = "I'm always optimistic about my future." All χ^2 tests and $\Delta\chi^2$ were significant, $p < .001$.

Table 7
Summary of Fit Indices From Comparative Factor Analysis (CFA) and Invariance Analyses Across Samples for the RS-10

Model—RS-10	χ^2 (*df*)	RMSEA [90% CI]	CFI	SRMR	ΔCFI	$\Delta\chi^2$
Single group CFA-original one factor model						
German sample	2850.419 (35)	.134 [.130, .139]	.902	.046		
Russian sample	723.443 (35)	.072 [.067, .077]	.924	.041		
Chinese sample	2935.914 (35)	.081 [.079, .084]	.892	.045		
Single group CFA ($\Theta_{1, 2;}$ $\Theta_{3, 7;}$ $\Theta_{8, 9}$ free)						
German sample	1503.380 (32)	.102 [.097, .106]	.949	.030		1347
Russian sample	362.651 (32)	.052 [.047, .057]	.963	.030		360.79
Chinese sample	2214.242 (32)	.074 [.071, .076]	.919	.039		721.67
Multiple group CFA models						
Configural	4080.273 (96)	.077 [.075, .079]	.938	.033		
Weak	5237.072 (114)	.081 [.079, .082]	.921	.056	.017	
Partial weak						
λ_3 free	4801.374 (112)	.078 [.076, .080]	.927	.051	.011	
$\lambda_{3,}$ λ_7 free	4650.362 (110)	.077 [.075, .079]	.930	.048	.008	
Strong	11083.724 (128)	.111 [.109, .113]	.830	.073	.1	
Partial strong						
τ_2 free	7784.767 (122)	.095 [.093, .097]	.881	.060	.049	
$\tau_{2,}$ τ_1 free	6591.783 (120)	.088 [.086, .090]	.900	.057	.030	
$\tau_{2,}$ $\tau_{1,}$ τ_4 free	5737.662 (118)	.083 [.081, .085]	.913	.053	.017	
$\tau_{2,}$ $\tau_{1,}$ $\tau_{4,}$ τ_6 free	5268.799 (116)	.080 [.078, .082]	.920	.051	.010	
$\tau_{2,}$ $\tau_{1,}$ $\tau_{4,}$ $\tau_{6,}$ τ_5 free	4941.333 (114)	.078 [.076, .080]	.925	.049	.005	

Note. RS-10 = Resilience Scale (10 items); RMSEA = root mean square error of analysis; CI = confidence interval; CFI = comparative fit index; SRMR = standardized root mean square residual; τ_2 = "I usually manage one way or another"; τ_1 = "When I make plans I follow through with them"; τ_4 = "I am friends with myself"; τ_5 = "I feel that I can handle many things at a time"; τ_6 = "I am determined." All χ^2 tests and $\Delta\chi^2$ were significant, $p < .001$.

pairs 1 ("When I make plans, I follow through with them") and 2 ("I usually manage one way or another"), 3 ("Keeping interested in things is important to me") and 7 ("I keep interested in things"), and 8 ("I can usually find something to laugh about") and 9 ("I can usually look at a situation in a number of ways"). Items 1 and 2 refer to the common theme of a certain goal-orientation/purposefulness. The content of Items 3 and 7 refers to a trait-like curiosity. Besides the very similar wording, Items 8 and 9 both refer to the ability to change one's own perspective. After allowing these error terms to correlate, the global model fit for the Russian and Chinese sample was acceptable, whereas the fit for the German sample was still poor. Modification indices indicated several more correlated error terms within the samples, but for the sake of comparability, these were not released in further analysis.

Measurement Invariance Across Cultures

Questionnaire-Social Support. Table 3 displays the model testing results of measurement invariance for the F-SozU K-14. The configural model yielded an acceptable global fit, implying that the configural invariance assumption holds. However, the RMSEA still indicated a poor fit. Imposing equality constraints on factor loadings led to a decrease in fit, but the ΔCFI was smaller than .01. Next, strong measurement invariance was tested. This yielded a poor fit to the data across groups, as ΔCFI exceeded .01. Therefore, partial measurement invariance was tested. Modification indices indicated a number of variant item intercepts across the samples. Consecutively, the equality constraints of intercepts of Item 3, 8, and 11 were released. The drop in the CFI was smaller than .01, and partial strong measurement invariance was established. However, the RMSEA still indicated several misspecifications, and global model fit was poor.

SHS. Table 4 displays the model testing results of measurement invariance of the SHS. In testing for measurement invariance, the configural invariance assumption for the SHS was supported, the model showed a good fit. Next, the factor loadings were constrained to be equal, and the drop in the CFI was larger than the proposed cut-off. Therefore, the model was tested for partial weak measurement invariance. By releasing the constraint of factor loadings of Items 3 and 4, the global model fit improved, with ΔCFI smaller than .01. The assumption of strong measurement invariance did not hold, as the drop in the CFI exceeded .01.

SWLS. Table 5 displays the model testing results of measurement invariance for the SWLS. The baseline model of the SWLS yielded an acceptable global fit, implying that the configural invariance assumption holds. By testing for weak measurement invariance, the drop in the CFI was larger than .01 compared to the baseline model. Modification indices indicated that the loading of Item 2 was not invariant. After releasing this constraint, partial weak measurement invariance was established. Partial strong measurement invariance was also established by releasing the equality constraints of Item 1 and Item 3. The global fit of this model remained acceptable.

Positive Mental Health Scale. Table 5 displays the model testing results of measurement invariance for the PMH-scale. As expected, all model comparisons conducted with chi-square difference testing were significant due to the large sample sizes. Therefore, the measurement invariance assumptions were tested taking the global model fit and ΔCFI into account. The model fit for the configural model yielded an acceptable fit and the model with invariant loadings still fit the data well. In addition, the assumption of strong measurement invariance was supported, al-

though fit indices were at the lower bound of the recommended cut-off criteria.

LOT-R. Table 6 displays the model testing results of measurement invariance for the LOT-R. The configural and weak model yielded a good fit to the data. The test for strong measurement invariance yielded a drop in the CFI that was too large compared to the weak measurement invariance model (ΔCFI = .024). Partial strong measurement invariance was established by allowing the intercept of Item 3 to vary.[1]

Resilience Scale. Table 7 displays the model testing results of measurement invariance for the RS-10. Although the fit for the German model for the RS-10 was poor, the scale was tested for measurement invariance. The global fit for the configural model was acceptable. In the weak measurement invariance model, item loadings were constrained to be equal. The global fit was not acceptable for this model, with ΔCFI larger than .01. The item loadings of Item 3 and Item 7 differed across the groups. In testing for strong measurement invariance, the drop in the CFI was also larger than .01. Modification indices indicated that the intercepts of items were invariant. After releasing the equality constraints in the descending order of Items 2, 1, 4, 6 and 5, partial strong measurement invariance was established.

Latent Mean Comparisons

Latent mean comparisons were conducted for the F-SozU K-14, SWLS, PMH-scale, and RS-10 (see Figure 1). The precondition for comparing latent means was not met by the SHS and the latent mean differences of the LOT-R could not be computed because of estimation difficulties. The latent mean calculation does not allow for conclusions regarding significant latent mean differences between Russian and Chinese students. To evaluate the significance of the latent mean differences between the Russian and Chinese samples, in a subsequent step the Russian sample was used as the reference group and latent mean differences were estimated. All mean differences between the Russian and Chinese samples were significant ($p < .001$). For the sake of parsimony, only the results for the latent mean comparisons in which the German mean is fixed to zero are reported. Comparison of the latent mean differences for the F-SozU K-14 indicated that compared to the German sample, the Russian and Chinese samples had significantly lower latent factor means ($z = -7.719$, $p < .001$, $d = 0.25$; $z = -18.540$, $p < .001$, $d = 0.35$). For the SWLS, the Russian and Chinese samples had significantly lower latent means than the German sample ($z = -6.357$, $p < .001$, $d = 0.20$; $z = -18.357$, $p < .001$, $d = 0.33$). The Russian and Chinese samples had significantly higher latent means on the PMH-scale than the German sample ($z = 8.559$, $p < .001$, $d = 0.28$; $z = 27.040$, $p < .001$, $d = 0.5$). For the RS-10, the Russian and Chinese samples had significantly higher latent means than the German sample ($z = 20.226$, $p < .001$, $d = 0.37$; $z = 11.226$, $p < .001$, $d = 0.37$).

Discussion

The aim of the present study was to investigate the cross-cultural measurement invariance of six positive constructs relevant to mental health and illness. The primary finding is that the factor structure of the scales of social support (F-SozU K-14), happiness (SHS), life satisfaction (SWLS), positive mental health (PMH-scale), and resilience (RS-10) are the same across German, Chinese, and Russian samples, as all demonstrated configural invariance. Only one questionnaire, the LOT-R, demonstrated an inconsistent correlation pattern and the model did not fit in the Chinese sample. Regarding the equality of loadings across cultures, full weak measurement invariance was shown for the F-SozU K-14 and PMH-scale across cultures. Full strong measurement invariance could only be established for the PMH-scale. All other scales, except the SHS, showed partial strong invariance across cultures. Latent means were compared for all questionnaires except for the SHS and the LOT-R. Results indicated significant differences in means across cultures, with German students experiencing significantly more social support and life satisfaction than Russian and Chinese students. In contrast, Chinese students reported higher positive mental health than German and Russian students. Regarding resilience, Russian students scored higher compared to German and Chinese students.

Configural and weak measurement invariance of the F-SozU K-14 was established, with acceptable global model fit. However, we did not find full strong measurement invariance. The intercepts of Item 3 ("I experience a lot of understanding and security from others"), Item 8 ("I have friends and family who simply just hug me"), and Item 11 ("There are people who share joy and sorrow with me") were higher in the Chinese sample than in the German and Russian samples. This means that Chinese students, given a certain level of social support, are less likely to endorse these items than German and Russian students. This finding is consistent with the latent factor mean comparison. German students scored significantly higher on the latent factor than Russian and Chinese students. However, the effects are small and this finding should not be overinterpreted. It might be due simply to the fact that the FSozU K-14 was developed in Germany, where a Western individualistic perspective might promote instrumental social support. Perhaps in more collectivistic cultures such as China, social support is less instrumental and more selflessly given (Markus & Kitayama, 1991). The universal construct of social support is probably not fully covered by the FSozU K-14.

The configural model of the SHS demonstrated a satisfactory fit to the data. In testing for full weak measurement invariance, Item 3 ("Some people are generally very happy. They enjoy life regardless of what is going on, getting the most out of everything") and Item 4 ("Some people are generally not very happy. Although they are not depressed, they never seem as happy as they might be") demonstrated variant loadings. The differing loadings of Items 3 and 4 could indicate either cultural or methodological differences. On a cultural level, the characterization of a happy and unhappy person can differ. Item 4 contrasts happiness with depression. However, it is unclear whether happiness is seen as the opposite of depression in all cultures, presupposing that happiness and depression are seen as psychological states. For example, the tendency of Chinese people to emphasize somatic symptoms of depression is widely acknowledged (Ryder, Yang, & Heine, 2002). A recent study showed that the Western tendency to put emphasis on

[1] The weak and strong measurement invariance models of the LOT-R did not converge in the first estimation procedure. Fit indices were calculated when standard errors and standard values were not computed. Therefore latent mean comparison cannot be conducted for this measure.

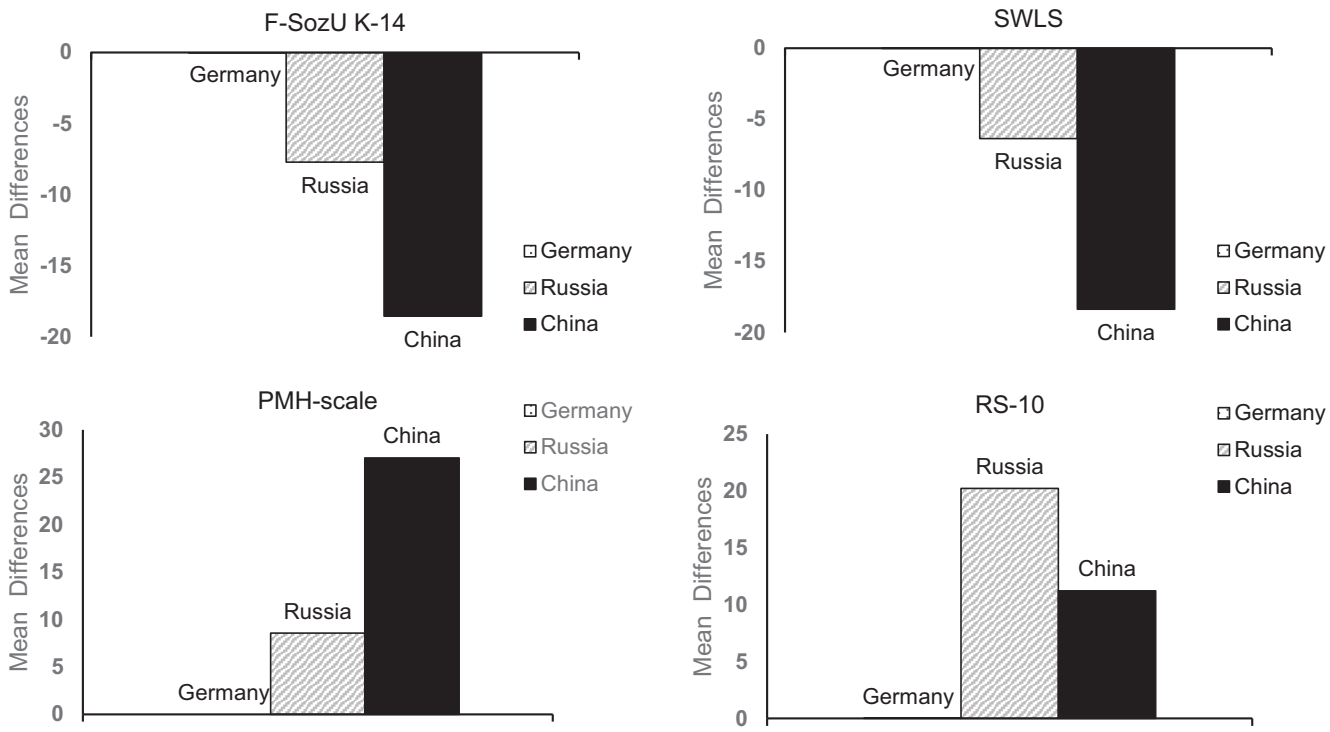

Figure 1. Latent factor means for the Russian and Chinese samples in comparison with the German sample (held at zero). The latent mean of the German student sample is fixed to zero. The differences are significant at $p < .001$. F-SozU K-14 = Questionnaire; Social Support; SHS = Subjective Happiness Scale; SWLS = Satisfaction with life Scale; PMH-scale = Positive Mental Health Scale; LOT-R = Life Orientation Test; RS-10 = Resilience Scale (10 items).

psychological symptoms may be more culturally specific than the Chinese emphasis on somatic symptoms (Ryder et al., 2008). From a methodological perspective, it may be that the reverse wording of Item 4 led to more confusion in participants than to the effect of response bias prevention (Van Sonderen, Sanderman, & Coyne, 2013).

The measurement invariance testing for the SWLS found large differences between cultures, in accordance with previous research (Oishi, 2006; Tucker et al., 2006; Zanon et al., 2014). At the level of factor loadings, Item 2 ("The conditions of my life are excellent") was variant across the samples. In the Chinese sample, this item is more closely related to the trait life satisfaction than in the German and Russian samples ($\lambda = .904$ compared to $\lambda = .667$ and $\lambda = .686$). It is difficult to interpret such across group differences since the SWLS claims to assess life satisfaction based on one's personal judgment and criteria determined by the person. Life conditions in Germany, Russia and China are different regarding, for example, socioeconomic or educational standards. It is not clear which standards German, Russian, and Chinese students refer to when they judge their life conditions. A possible explanation could be that Chinese students place more value on the opportunity to study at a university than German and Russian students. At the level of intercepts, Item 1 ("In most ways my life is close to my ideal") and Item 3 ("I am satisfied with my life") also showed noninvariance, with German and Russian students more likely to endorse these items than Chinese students. Asian students experience more academic stress than Western students (Ang et al., 2009; Lee & Larson, 2000; Liu & Lu, 2012), resulting in lower satisfaction with life. The latent mean comparison indicates that German students have a significantly higher latent mean on the remaining two invariant items representing a higher life satisfaction than Russian and Chinese students; however, the effects are small.

For the PMH-scale, full strong measurement invariance was established. Russian and Chinese students had significantly higher means on the latent factor positive mental health than German students, with a small effect for the Russian sample, and a medium effect for the Chinese sample. Taking the previous results of the latent mean comparison into account, this finding is surprising in indicating that Chinese students experience less social support and report lower life satisfaction than German and Russian students. The content of the PMH-scale items is very similar to items referring to the construct peace of mind (PoM; Lee, Lin, Huang & Frederickson, 2013), such as that of Item 9 ("I am a calm, balanced human being"). This construct is defined as an internal state of peacefulness and harmony and represents the low-arousal positive affect valued in the Chinese culture (Lee et al., 2013). Further studies should examine the construct validity of PoM, especially in Western cultures, and assess the convergent validity of the PMH-scale with PoM. A methodological limitation concerns the evaluation of models estimated with WLSMV by ΔCFI. Some studies have indicated that the ΔCFI criterion should be limited to models

estimated by the maximum likelihood method (Cheung & Rensvold, 2002). However, recent work has shown that this criterion is still the best alternative to identify measurement invariance in large sample sizes (Sass et al., 2014). Sass and colleagues (2014) also recommend comparing different estimation procedures and cutoff criteria. Therefore, we used ML to estimate the model parameters. We found that the single models of the PMH-scale estimated with ML did not fit the data adequately. In contrast, the single models of the PMH-scale with WLSMV estimator outperformed the ML estimator and seemed to be well specified.

As a consequence of the gross misfit at the level of item-correlations in the Chinese sample, analyses for the LOT-R were conducted only in the German and Russian samples. Similarly to Monzani et al. (2014), we found that responses were best modeled by one trait-factor assessing optimism and one response-style factor for the three positively worded items. Regarding measurement invariance testing, only the intercepts of Item 3 were variant across groups. Thus, the LOT-R proved to be partial strong measurement invariant. We can only speculate as to why this scale did not work for the Chinese sample. Our results are consistent with the findings by Lai and Yue (2000) who were able to apply the unidimensional LOT-R in a Chinese sample from Hong Kong but not in a mainland Chinese sample. The authors assume that the Western conceptualization of optimism is not fully congruent with the Chinese concept of optimism. Lai adapted and revised the Chinese LOT-R in more recent studies (Chan, Lai, & Wong, 2006; Lai, 2009). Our translation of Item 3 ("If something can go wrong for me, it will") fits well the English language version but differs from Item 3 translated by Lai. Also our version of Item 7 ("I hardly ever expect things to go my way") differs slightly from the Chinese version. However, in our opinion the core message of this item remains the same. Further analyses should reinvestigate the construct of optimism in Asian cultures to ensure its universal applicability or take cultural specificities into account and use the Chinese LOT-R version adapted by Lai.

The RS-10 proved to be partially strong invariant. Partial weak measurement invariance was established by releasing the equality constraints of Items 3 and 7. These items were more strongly related to the trait resilience in the German sample than in the Russian and Chinese samples. Partial strong measurement invariance and an acceptable fit was achieved by releasing the intercepts of Items 1 ("When I make plans I follow through with them"), 2 ("I usually manage one way or another"), 4 ("I am friends with myself"), 5 ("I feel that I can handle many things at a time"), and 6 ("I am determined"). The pattern of variant items across groups was again mixed. The interpretation of the latent mean differences should be made with caution, as the comparisons are based on three items. Russian and Chinese students score significantly higher on the latent trait resilience than German students, with small effects. One explanation for the unacceptable fit of the RS-11 in the German sample could lie in the item selection process for the RS-11, which was based on statistical criteria rather on theoretical considerations (Schumacher et al., 2005). Item 10 did not function well in the Russian and Chinese sample. A reexamination of the wordings across cultures indicated that the translation was not fully appropriate in the Russian and Chinese samples. Our findings concerning measurement invariance testing are in line with the results of Yang et al. (2012) who found that only four of 14 items were invariant among U.S., Taiwanese, and Chinese students. Cross-cultural comparisons should be made with caution. It can be assumed that *resilience*, defined as a multidimensional construct, has culture-specific aspects that should be examined and considered more deeply in further cross-cultural research.

This study has several limitations that need to be taken into account. First, data in Germany were collected at one large university, whereas data in Russia and China were collected at three and five different universities, respectively, which were either public or private. Such heterogeneity in Russia and China could have influenced the data with results mediated by socioeconomic status. This could be tested in further studies. Second, the reasons for nonequivalence still remain unclear. Cultural and methodological bias could not be disentangled because there are many different factors that could have contributed to the nonequivalence of items (van de Vijver, 2007). Therefore, the interpretation of results is not conclusive and the source of bias can only be hypothesized. Further studies are needed to examine construct or methodological bias, as the examination of measurement invariance can only detect item biases. Investigating social desirability in a cross-cultural framework and controlling for its effects may be a possible next step. Third, given the scales fulfill the assumption of partial strong measurement invariance, latent means can be compared, however, only the unbiased items are included in the comparisons. This can lead to a construct being underrepresented and therefore to shortcomings in the interpretation of cross-cultural comparisons. Fourth, the analyses were conducted in student samples, and thus cannot necessarily be generalized to the population level. It can be hypothesized that measurement invariance is more difficult to establish across representative or community adult samples, given the greater heterogeneity of such samples. With the increasing globalization process, student samples are likely to share more common values and norms than older community-based samples.

In summary, we find that most positive constructs can be considered universal across the nations included in this study, even though individual items show cultural differences. The results of this study are in line with previous studies examining positive constructs cross-culturally (Tucker et al., 2006; Yang et al., 2012). Full strong measurement invariance was difficult to establish, perhaps in some cases due to methodological artifacts, such as translation problems for certain scales, but may also be due to culture-specific expressions of these constructs. This study has also demonstrated that observed mean differences among German, Russian, and Chinese students cannot be fully explained by mean differences in latent factors. Caution is necessary when interpreting possible cross-cultural differences assessed by observed mean comparisons, because measurement artifacts can mask true group differences. On a more general level, this study should raise the awareness of researchers aiming to conduct cross-cultural comparisons, especially when comparing observed means in the field of clinical, social, health, and positive psychology. Group differences on a given construct can be due to instrument bias across cultures and thus may not truly reflect differences in the latent construct, potentially leading to premature or inaccurate conclusions in cross-cultural research. Future research should take culture-specific expressions into account to achieve unbiased measures. In addition, underrepresentation of psychological constructs should be avoided or at least explicitly stated. Measurement sensitivity to cross-cultural issues can be promoted through greater international ac-

ademic cooperation and professional exchange among cross-cultural researchers.

References

Alessandri, G., Vecchione, M., Fagnani, C., Bentler, P. M., Barbaranelli, C., Medda, E., . . . Caprara, G. V. (2010). Much more than model fitting? Evidence for the heritability of method effect associated with positively worded items of the Life Orientation Test Revised. *Structural Equation Modeling, 17,* 642–653. http://dx.doi.org/10.1080/10705511.2010.510064

Andersson, G. (1996). The benefits of optimism: A meta-analytic review of the Life Orientation Test. *Personality and Individual Differences, 21,* 719–725. http://dx.doi.org/10.1016/0191-8869(96)00118-3

Ang, R. P., Klassen, R. M., Chong, W. H., Huan, V. S., Wong, I. Y., Yeo, L. S., & Krawchuk, L. L. (2009). Cross-cultural invariance of the Academic Expectations Stress Inventory: Adolescent samples from Canada and Singapore. *Journal of Adolescence, 32,* 1225–1237. http://dx.doi.org/10.1016/j.adolescence.2009.01.009

Baas, K. D., Cramer, A. O., Koeter, M. W., van de Lisdonk, E. H., van Weert, H. C., & Schene, A. H. (2011). Measurement invariance with respect to ethnicity of the Patient Health Questionnaire-9 (PHQ-9). *Journal of Affective Disorders, 129,* 229–235. http://dx.doi.org/10.1016/j.jad.2010.08.026

Beauducel, A., & Herzberg, P. Y. (2006). On the performance of maximum likelihood versus means and variance adjusted weighted least squares estimation in CFA. *Structural Equation Modeling, 13,* 186–203. http://dx.doi.org/10.1207/s15328007sem1302_2

Becker, P. (1989). *Der Trier Persönlichkeitsfragebogen (TPF) Handanweisung* [The Trier Personality Inventory (TIP) Manual]. Göttingen, Germany: Hogrefe.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107,* 238–246. http://dx.doi.org/10.1037/0033-2909.107.2.238

Borsboom, D. (2006). When does measurement invariance matter? *Medical Care, 44*(Suppl. 3), S176–S181. http://dx.doi.org/10.1097/01.mlr.0000245143.08679.cc

Bradburn, N. M. (1969). *The structure of psychological well-being*. Oxford, England: Aldine.

Brislin, R. W. (1970). Back-translation for cross-cultural research. *Journal of Cross-Cultural Psychology, 1,* 185–216. http://dx.doi.org/10.1177/135910457000100301

Brunwasser, S. M., Gillham, J. E., & Kim, E. S. (2009). A meta-analytic review of the Penn Resiliency Program's effect on depressive symptoms. *Journal of Consulting and Clinical Psychology, 77,* 1042–1054. http://dx.doi.org/10.1037/a0017671

Byrne, B. M., Shavelson, R. J., & Muthén, B. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin, 105,* 456–466. http://dx.doi.org/10.1037/0033-2909.105.3.456

Carver, C. S., Scheier, M. F., & Segerstrom, S. C. (2010). Optimism. *Clinical Psychology Review, 30,* 879–889. http://dx.doi.org/10.1016/j.cpr.2010.01.006

Chan, I. W., Lai, J. C., & Wong, K. W. (2006). Resilience is associated with better recovery in Chinese people diagnosed with coronary heart disease. *Psychology & Health, 21,* 335–349. http://dx.doi.org/10.1080/14768320500215137

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9,* 233–255. http://dx.doi.org/10.1207/S15328007SEM0902_5

Cohen, S. (2004). Social relationships and health. *American Psychologist, 59,* 676–684.

Crockett, L. J., Randall, B. A., Shen, Y. L., Russell, S. T., & Driscoll, A. K. (2005). Measurement equivalence of the center for epidemiological studies depression scale for Latino and Anglo adolescents: A national study. *Journal of Consulting and Clinical Psychology, 73,* 47–58. http://dx.doi.org/10.1037/0022-006X.73.1.47

Dere, J., Watters, C. A., Yu, S. C. M., Bagby, R. M., Ryder, A. G., & Harkness, K. L. (2015). Cross-cultural examination of measurement invariance of the Beck Depression Inventory-II. *Psychological Assessment, 27,* 68–81. http://dx.doi.org/10.1037/pas0000026

Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The Satisfaction With Life Scale. *Journal of Personality Assessment, 49,* 71–75. http://dx.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. http://dx.doi.org/10.1037/0033-2909.125.2.276

Extremera, N., & Fernández-Berrocal, P. (2014). The Subjective Happiness Scale: Translation and preliminary psychometric evaluation of a Spanish version. *Social Indicators Research, 119,* 473–481. http://dx.doi.org/10.1007/s11205-013-0497-2

Fahrenberg, J., Selg, H., & Hampel, R. (1989). *Das Freiburger Persönlichkeitsinventar: FPI; Handanweisung; revidierte Fassung FPI-R und teilweise geänderte Fassung FPI-A1* [The Freiburg Personality Inventory: FPI; Manual; revised version FPI-R and partly adapted version FPI-A1]. Göttingen, Germany: Verlag für Psychologie Hogrefe.

Fletcher, D., & Sarkar, M. (2013). Psychological resilience: A review and critique of definitions, concepts, and theory. *European Psychologist, 18,* 12–23. http://dx.doi.org/10.1027/1016-9040/a000124

Ford, J. K., MacCallum, R. C., & Tait, M. (1986). The application of exploratory factor analysis in applied psychology: A critical review and analysis. *Personnel Psychology, 39,* 291–314. http://dx.doi.org/10.1111/j.1744-6570.1986.tb00583.x

Fydrich, T., Sommer, G., Menzel, U., & Höll, B. (1987). The Social Support Questionnaire (short form; F-SozU-K-22). *Zeitschrift für klinische Psychologie und Psychotherapie, 16,* 434–436.

Fydrich, T., Sommer, G., Tydecks, S., & Brähler, E. (2009). Fragebogen zur sozialen Unterstützung (F-SozU): Normierung der Kurzform (K-14). [Social Support Questionnaire (F-SozU): Standardization of short form (K-14)]. *Zeitschrift für Medizinische Psychologie, 18,* 43–48.

Gallagher, M. W., Lopez, S. J., & Pressman, S. D. (2013). Optimism is universal: Exploring the presence and benefits of optimism in a representative sample of the world. *Journal of Personality, 81,* 429–440. http://dx.doi.org/10.1111/jopy.12026

Glaesmer, H., Grande, G., Braehler, E., & Roth, M. (2011). The German version of the Satisfaction with Life Scale (SWLS): Psychometric properties, validity, and population-based norms. *European Journal of Psychological Assessment, 27,* 127–132. http://dx.doi.org/10.1027/1015-5759/a000058

Glaesmer, H., Hoyer, J., Klotsche, J., & Herzberg, P. Y. (2008). Die deutsche Version des Life-Orientation-Tests (LOT-R) zum dispositionellen Optimismus und Pessimismus [The German Version of the Life-Orientation-Test (LOT-R) for dispositional optimism and pessimism]. *Zeitschrift für Gesundheitspsychologie, 16,* 26–31. http://dx.doi.org/10.1026/0943-8149.16.1.26

Grob, A., Lüthi, R., Kaiser, F. G., Flammer, A., Mackinnon, A., & Wearing, A. J. (1991). Berner Fragebogen zum Wohlbefinden Jugendlicher (BFW) [The Bern Subjective Well-Being Questionnaire for Adolescents (BFW)]. *Diagnostica, 37,* 66–75.

Guerra, M., Ferri, C., Llibre, J., Prina, A. M., & Prince, M. (2015). Psychometric properties of EURO-D, a geriatric depression scale: A cross-cultural validation study. *BMC Psychiatry, 15,* 12. http://dx.doi.org/10.1186/s12888-015-0390-4

Herzberg, P. Y., Glaesmer, H., & Hoyer, J. (2006). Separating optimism and pessimism: A robust psychometric analysis of the revised Life Orientation Test (LOT-R). *Psychological Assessment, 18,* 433–438. http://dx.doi.org/10.1037/1040-3590.18.4.433

Hirschfeld, G., & von Brachel, R. (2014). Multiple-group confirmatory factor analysis in R–A tutorial in measurement invariance with contin-

uous and ordinal indicators. *Practical Assessment, Research & Evaluation, 19,* 2. Retrieved from http://pareonline.net/pdf/v19n7.pdf

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

Iani, L., Lauriola, M., Layous, K., & Sirigatti, S. (2014). Happiness in Italy: Translation, factorial structure and norming of the Subjective Happiness Scale in a large community sample. *Social Indicators Research, 118,* 953–967. http://dx.doi.org/10.1007/s11205-013-0468-7

Jöreskog, K. G. (1971). Simultaneous factor analysis in several populations. *Psychometrika, 36,* 409–426. http://dx.doi.org/10.1007/BF02291366

Kendler, K. S., Aggen, S. H., Li, Y., Lewis, C. M., Breen, G., Boomsma, D. I., . . . Flint, J. (2015). The similarity of the structure of *DSM–IV* criteria for major depression in depressed women from China, the United States and Europe. *Psychological Medicine, 45,* 1945–1954. http://dx.doi.org/10.1017/S0033291714003067

Keyes, C. L. (2002). The mental health continuum: From languishing to flourishing in life. *Journal of Health and Social Behavior, 43,* 207–222. http://dx.doi.org/10.2307/3090197

Kline, R. B. (2010). *Principles and practice of structural equation modeling* (3rd ed.). New York, NY: Guilford Press.

Lai, J. C. (2009). Dispositional optimism buffers the impact of daily hassles on mental health in Chinese adolescents. *Personality and Individual Differences, 47,* 247–249. http://dx.doi.org/10.1016/j.paid.2009.03.007

Lai, J. C., & Yue, X. (2000). Measuring optimism in Hong Kong and mainland Chinese with the revised Life Orientation Test. *Personality and Individual Differences, 28,* 781–796. http://dx.doi.org/10.1016/S0191-8869(99)00138-5

Lee, M., & Larson, R. (2000). The Korean "examination hell": Long hours of studying, distress, and depression. *Journal of Youth and Adolescence, 29,* 249–271. http://dx.doi.org/10.1023/A:1005160717081

Lee, Y.-C., Lin, Y.-C., Huang, C.-L., & Fredrickson, B. L. (2013). The construct and measurement of peace of mind. *Journal of Happiness Studies, 14,* 571–590. http://dx.doi.org/10.1007/s10902-012-9343-5

Lindorff, M. (2000). Is it better to perceive than receive? Social support, stress and strain for managers. *Psychology, Health & Medicine, 5,* 271–286. http://dx.doi.org/10.1080/713690199

Liu, Y., & Lu, Z. (2012). Chinese high school students' academic stress and depressive symptoms: Gender and school climate as moderators. *Stress and Health: Journal of the International Society for the Investigation of Stress, 28,* 340–346. http://dx.doi.org/10.1002/smi.2418

Losada, A., de los Angeles Villareal, M., Nuevo, R., Márquez-González, M., Salazar, B. C., Romero-Moreno, R., . . . Fernáñdez-Fernández, V. (2012). Cross-cultural confirmatory factor analysis of the CES-D in Spanish and Mexican dementia caregivers. *The Spanish Journal of Psychology, 15,* 783–792. http://dx.doi.org/10.5209/rev_SJOP.2012.v15.n2.38890

Lovibond, P. F., & Lovibond, S. H. (1995). The structure of negative emotional states: Comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. *Behaviour Research and Therapy, 33,* 335–343. http://dx.doi.org/10.1016/0005-7967(94)00075-U

Lukat, J., Margraf, J., Lutz, R., van der Veld, W. M., & Becker, E. S. (2016). Psychometric properties of the Positive Mental Health Scale (PMH-scale). *BMC Psychology, 4,* 8. http://dx.doi.org/10.1186/s40359-016-0111-x

Lyubomirsky, S., & Lepper, H. S. (1999). A measure of subjective happiness: Preliminary reliability and construct validation. *Social Indicators Research, 46,* 137–155. http://dx.doi.org/10.1023/A:1006824100041

Lyubomirsky, S., Tkach, C., & DiMatteo, M. R. (2006). What are the differences between happiness and self-esteem? *Social Indicators Research, 78,* 363–404. http://dx.doi.org/10.1007/s11205-005-0213-y

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods, 1,* 130–149. http://dx.doi.org/10.1037/1082-989X.1.2.130

Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review, 98,* 224–253. http://dx.doi.org/10.1037/0033-295X.98.2.224

Marum, G., Clench-Aas, J., Nes, R. B., & Raanaas, R. K. (2014). The relationship between negative life events, psychological distress and life satisfaction: A population-based study. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care & Rehabilitation, 23,* 601–611. http://dx.doi.org/10.1007/s11136-013-0512-8

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika, 58,* 525–543. http://dx.doi.org/10.1007/BF02294825

Monzani, D., Steca, P., & Greco, A. (2014). Brief report: Assessing dispositional optimism in adolescence—factor structure and concurrent validity of the Life Orientation Test—Revised. *Journal of Adolescence, 37,* 97–101. http://dx.doi.org/10.1016/j.adolescence.2013.11.006

Muthén, B., & Asparouhov, T. (2002). Latent variable analysis with categorical outcomes: Multiple-group and growth modeling in Mplus. *Mplus Web Notes, 4,* 1–22. Retrieved from http://www.statmodel.com/examples/webnote.shtml

Nover, T. (2012). *A cross-cultural comparison of mental health status in adolescents* (Doctoral dissertation). Retrieved from http://ubt.opus.hbz-nrw.de/volltexte/2012/778

Nuevo, R., Dunn, G., Dowrick, C., Vázquez-Barquero, J. L., Casey, P., Dalgard, O. S., . . . Ayuso-Mateos, J. L. (2009). Cross-cultural equivalence of the Beck Depression Inventory: A five-country analysis from the ODIN study. *Journal of Affective Disorders, 114,* 156–162. http://dx.doi.org/10.1016/j.jad.2008.06.021

Oishi, S. (2006). The concept of life satisfaction across cultures: An IRT analysis. *Journal of Research in Personality, 40,* 411–423. http://dx.doi.org/10.1016/j.jrp.2005.02.002

Oishi, S. (2007). The application of structural equation modeling and item response theory to cross-cultural positive psychology research. In A. D. Ong & M. H. M. van Dulmen (Eds.), *Oxford handbook of methods in positive psychology* (pp. 126–138). New York, NY: Oxford University Press.

Pavot, W., & Diener, E. (2008). The satisfaction with life scale and the emerging construct of life satisfaction. *The Journal of Positive Psychology, 3,* 137–152. http://dx.doi.org/10.1080/17439760701756946

Rasmussen, H. N., Scheier, M. F., & Greenhouse, J. B. (2009). Optimism and physical health: A meta-analytic review. *Annals of Behavioral Medicine, 37,* 239–256. http://dx.doi.org/10.1007/s12160-009-9111-x

Rauch, W. A., Schweizer, K., & Moosbrugger, H. (2007). Method effects due to social desirability as a parsimonious explanation of the deviation from unidimensionality in LOT-R scores. *Personality and Individual Differences, 42,* 1597–1607. http://dx.doi.org/10.1016/j.paid.2006.10.035

Renshaw, T., & Cohen, A. (2014). Life satisfaction as a distinguishing indicator of college student functioning: Further validation of the two-continua model of mental health. *Social Indicators Research, 117,* 319–334. http://dx.doi.org/10.1007/s11205-013-0342-7

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

Ryder, A. G., Yang, J., & Heine, S. J. (2002). Somatization vs. psychologization of emotional distress: A paradigmatic example for cultural

psychopathology. *Online Readings in Psychology and Culture, 10,* 3. http://dx.doi.org/10.9707/2307-0919.1080

Ryder, A. G., Yang, J., Zhu, X., Yao, S., Yi, J., Heine, S. J., & Bagby, R. M. (2008). The cultural shaping of depression: Somatic symptoms in China, psychological symptoms in North America? *Journal of Abnormal Psychology, 117,* 300–313. http://dx.doi.org/10.1037/0021-843X.117.2.300

Sass, D. A., Schmitt, T. A., & Marsh, H. W. (2014). Evaluating model fit with ordered categorical data within a measurement invariance framework: A comparison of estimators. *Structural Equation Modeling, 21,* 167–180. http://dx.doi.org/10.1080/10705511.2014.882658

Scheier, M. F., & Carver, C. S. (1985). Optimism, coping, and health: Assessment and implications of generalized outcome expectancies. *Health Psychology, 4,* 219–247. http://dx.doi.org/10.1037/0278-6133.4.3.219

Scheier, M. F., Carver, C. S., & Bridges, M. W. (1994). Distinguishing optimism from neuroticism (and trait anxiety, self-mastery, and self-esteem): A reevaluation of the Life Orientation Test. *Journal of Personality and Social Psychology, 67,* 1063–1078. http://dx.doi.org/10.1037/0022-3514.67.6.1063

Schumacher, J., Leppert, K., Gunzelmann, T., Strauss, B., & Brähler, E. (2005). Die Resilienzskala-Ein Fragebogen zur Erfassung der psychischen Widerstandsfähigkeit als Personmerkmal [The Resilience Scale-A questionnaire for the assessment of psychosocial stress-resistance as a personality factor]. *Zeitschrift für Klinische Psychologie, Psychiatrie und Psychotherapie, 53,* 16–39.

Skriner, L. C., & Chu, B. C. (2014). Cross-ethnic measurement invariance of the SCARED and CES-D in a youth sample. *Psychological Assessment, 26,* 332–337. http://dx.doi.org/10.1037/a0035092

Steenkamp, J.-B. E. M., & Baumgartner, H. (1998). Assessing measurement invariance in cross-national consumer research. *The Journal of Consumer Research, 25,* 78–90. http://dx.doi.org/10.1086/209528

Steiger, J. H. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioral Research, 25,* 173–180. http://dx.doi.org/10.1207/s15327906mbr2502_4

Swami, V., Stieger, S., Voracek, M., Dressler, S. G., Eisma, L., & Furnham, A. (2009). Psychometric evaluation of the Tagalog and German Subjective Happiness Scales and a cross-cultural comparison. *Social Indicators Research, 93,* 393–406. http://dx.doi.org/10.1007/s11205-008-9331-7

Tönnies, S., Plöhn, S., & Krippendorf, U. (1996). *Skalen zur psychischen Gesundheit (SPG)* [Scales for mental health (SPG)]. Heidelberg, Germany: Asanger.

Tucker, K. L., Ozer, D. J., Lyubomirsky, S., & Boehm, J. K. (2006). Testing for measurement invariance in the satisfaction with life scale: A comparison of Russians and North Americans. *Social Indicators Research, 78,* 341–360. http://dx.doi.org/10.1007/s11205-005-1037-5

Uchino, B. N., Cacioppo, J. T., & Kiecolt-Glaser, J. K. (1996). The relationship between social support and physiological processes: A review with emphasis on underlying mechanisms and implications for health. *Psychological Bulletin, 119,* 488–531. http://dx.doi.org/10.1037/0033-2909.119.3.488

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. http://dx.doi.org/10.1177/109442810031002

van De Vijver, F. J. R. (2007). Cultural and gender differences in gender-role beliefs, sharing household task and child-care responsibilities, and well-being among immigrants and majority members in the Netherlands. *Sex Roles, 57,* 813–824. http://dx.doi.org/10.1007/s11199-007-9316-z

van Sonderen, E., Sanderman, R., & Coyne, J. C. (2013). Ineffectiveness of reverse wording of questionnaire items: Let's learn from cows in the rain. *PLoS ONE, 8,* e68967. http://dx.doi.org/10.1371/journal.pone.0068967

Veenhoven, R. (1996). Developments in satisfaction-research. *Social Indicators Research, 37,* 1–46. http://dx.doi.org/10.1007/BF00300268

Wagnild, G. M., & Young, H. M. (1993). Development and psychometric evaluation of the Resilience Scale. *Journal of Nursing Measurement, 1,* 165–178. http://www.resiliencescale.com/wp-content/uploads/2014/06

Whisman, M. A., Judd, C. M., Whiteford, N. T., & Gelhorn, H. L. (2013). Measurement invariance of the Beck Depression Inventory-2nd ed. (BDI-II) across gender, race, and ethnicity in college students. *Assessment, 20,* 419–428. http://dx.doi.org/10.1177/1073191112460273

Wissing, M. P., & Temane, Q. M. (2008). The structure of psychological well-being in cultural context: Towards a hierarchical model of psychological health. *Journal of Psychology in Africa, 18,* 45–55.

Wood, A. M., & Tarrier, N. (2010). Positive clinical psychology: A new vision and strategy for integrated research and practice. *Clinical Psychology Review, 30,* 819–829. http://dx.doi.org/10.1016/j.cpr.2010.06.003

Wu, W., Lu, Y., Tan, F., Yao, S., Steca, P., Abela, J. R., & Hankin, B. L. (2012). Assessing measurement invariance of the children's depression inventory in Chinese and Italian primary school student samples. *Assessment, 19,* 506–516. http://dx.doi.org/10.1177/1073191111421286

Yang, Y., Li, M.-H., & Xia, Y. (2012). Measurement invariance of the Resilience Scale. *The International Journal of Educational and Psychological Assessment, 11,* 1–19. Retrieved from https://sites.google.com/site/tijepa2012/vol-11-2

Zanon, C., Bardagi, M., Layous, K., & Hutz, C. (2014). Validation of the Satisfaction with Life Scale to Brazilians: Evidences of Measurement Noninvariance Across Brazil and US. *Social Indicators Research, 119,* 443–453. http://dx.doi.org/10.1007/s11205-013-0478-5

Zhang, B., Fokkema, M., Cuijpers, P., Li, J., Smits, N., & Beekman, A. (2011). Measurement invariance of the center for epidemiological studies depression scale (CES-D) among chinese and dutch elderly. *BMC Medical Research Methodology, 11,* 74.

Received March 16, 2015
Revision received May 11, 2016
Accepted May 19, 2016 ■