Support for a general factor of well-being

Ylenio Longo, Iain Coyne and Stephen Joseph

University of Nottingham

Petter Gustavsson

Karolinska Institutet

Longo, Y., Coyne, I., Joseph, S., & Gustavsson, P. (in press). Support for a general factor of well-being. *Personality and Individual Differences*. doi: 10.1016/j.paid.2016.03.082

Author note

Ylenio Longo, Division of Psychiatry and Applied Psychology, University of Nottingham, UK; Iain Coyne, Division of Psychiatry and Applied Psychology, University of Nottingham, UK; Stephen Joseph, School of Education, University of Nottingham, UK; Petter Gustavsson, Department of Clinical Neuroscience, Karolinska Institutet, Sweden.

This research was supported by the Economic and Social Research Council grant ES/J500100/1 to Ylenio Longo. Portions of this article served as part of Ylenio Longo's PhD thesis.

Correspondence concerning this article should be addressed to Ylenio Longo, School of Medicine, Division of Psychiatry and Applied Psychology, University of Nottingham, YANG Fujia Building, Jubilee Campus, Nottingham NG8 1BB. E-mail: yleniolongo@gmail.com.

Abstract. Well-being is typically defined as positive feeling (e.g. happiness), positive functioning (e.g. competence, meaning) or a combination of the two. Recent evidence indicates that well-being indicators belonging to different categories can be explained by single "general" factor of well-being (e.g. Jovanovic, 2015). We further test this hypothesis using a recent well-being scale, which includes indicators of positive feeling and positive functioning (Huppert & So, 2013). While the authors of the scale originally identified a two-factor structure, in view of recent evidence, we hypothesize that the two-factor solution may be due to a method effect of different items being measured with different rating scales. In study 1, we use data from the European Social Survey round 3 (*n* = 41,461) and find that two factors have poor discriminant validity and, after using a bifactor model to account for different rating scales, only the general factor is reliable. In study 2, we eliminate method effects by using the same rating scale across items, recruit a new sample (*n* = 507), and find that a one-factor model fits the data well. The results support the hypothesis that well-being indicators, typically categorized as "positive feeling" and "positive functioning," reflect a single general factor.

Keywords. structure of well-being; flourishing; factor analysis; unidimensionality; positive feeling; positive functioning; hedonic well-being; eudaimonic well-being

1. Introduction

Well-being is often defined as a set of positive feeling (e.g. happiness, satisfaction) and positive functioning elements (e.g. engagement, self-acceptance) (Ryan & Deci, 2001). Some operationalizations of well-being focus on positive feeling (e.g. Diener, 1984), some focus on positive functioning (e.g. Ryff, 1989), while others include elements of both (e.g. Huppert & So, 2013; Keyes, 2002).

Despite the conceptual distinction between positive feeling and positive functioning, the two have been found to be highly correlated, with relationships as high as .76 and .84 reported (Linley, Maltby, Wood, Osborne & Hurling, 2009; Keyes, Shmotkin & Ryff, 2002). Furthermore, recent evidence suggests that well-being indicators, which are typically categorized as positive feeling and positive functioning, may be explained by a single higher-order or "general" factor. For example, Gallagher, Lopez and Preacher (2009) tested several models in a student sample (n = 591) and a representative U.S. sample (n = 4,032), and found that a single higher-order factor adequately explained the relationships among different well-being indicators. More complex models, with two or more higher-order factors, fit the data slightly better, at the expense of parsimony. These findings were corroborated in a large international sample (n = 7,617) showing that positive feeling and positive functioning factors had a latent correlation of .96, exhibited similar correlations with external criteria, and their components fitted a single higher-order factor (Disabato, Goodman, Kashdan, Short & Jarden, in press).

This hypothesis has also been tested using a bifactor model. A bifactor model assumes that each indicator may reflect both a general factor (e.g. well-being) and specific factors (e.g. positive feeling or positive functioning). Thus, in a bifactor model each indicator is allowed to load both on a general factor and on a specific factor. In several studies a bifactor model fit well-being data better than a higher-order model (Chen, Jing, Hayes & Lee, 2013; Jovanovic, 2015) with the general factor explaining most of the variance in well-being indicators (Chen et al., 2013; de Bruin & Du Plessis, 2015; Jovanovic, 2015). These findings suggest that indicators of well-being (e.g. happiness, self-acceptance) may be adequately explained by a single higher-order or general factor.

In this article, we aim to build on these findings using a recently developed operationalization and measure of well-being, which includes elements of positive feeling and positive functioning (Huppert & So, 2013). The operationalization includes ten indicators: happiness, emotional stability, vitality, resilience, optimism, engagement, competence, meaning, positive relationships and self-esteem. These indicators were identified by defining the opposites of symptoms of depression and anxiety. This operationalization makes an important contribution to research, as it combines a comprehensive list of well-being indicators and shows how these may be linked to symptoms of psychological health problems.

Based on their operationalization, Huppert and So (2013) identified ten items, measuring each well-being indicator, from the European Social Survey Round 3. After factor analyzing the data, they found that the ten items loaded on two separate factors. The wording of all items but one was positive. However, the rating scales were scored in two different directions. Four items, measuring happiness, emotional stability, vitality and resilience, were scored from low to high: low ratings indicating low levels of the construct and high rating indicating high levels of the construct. Conversely, the other six items, measuring optimism, engagement, competence, meaning, positive relationships and self-esteem, were scored from high to low: low ratings indicating high levels of the construct and high ratings indicating low levels of the construct.

Each of the two factors that emerged in Huppert and So's (2013) study was almost exclusively made up of items scored in the same direction. Specifically, four items scored from low to high loaded on the first factor[1], four items scored from high to low loaded on the second factor, and two items scored from high to low loaded on the first factor, but exhibited some cross-loadings on the second factor (e.g. .28).

When items are scored in opposite directions, they often produce separate factors in a factor analysis. For example, two simulation studies have shown that, if only 10% of respondents complete a questionnaire carelessly, thus not noticing the change in rating scale, factor analytic results will not support a one-factor solution even though the construct measured is unidimensional (Schmitt & Stults, 1985; Woods, 2006). Instead, the analyses will indicate that a second factor or a method factor is necessary to account for the negatively-scored items. However, these factors would be artifacts due to the measurement method, rather than substantial differences in meaning among the items.

Two corrections are often used to account for these method effects. First, one can specify that the errors among negatively- or positively-scored items of a construct correlate. These correlated errors reduce the methodological bias that produced different responses to positive and negative items (Bachman & O'Malley, 1986; Marsh, 1996). Alternatively, one can use a bifactor model by specifying a general factor and two "method" factors, explaining the variance in positively and negatively scored items (e.g. Alessandri, Vecchione, Eisenberg & Laguna, 2015). One advantage of a bifactor model over a correlated errors model is that the magnitude of the method effects is more readily interpretable (Brown, 2015), as one can inspect the factor loadings and reliability of each factor.

Based on the evidence outlined hitherto, we hypothesized that after controlling for methods effects, Huppert and So's (2013) well-being scale would exhibit a strong general factor. We tested this hypothesis in two studies.

[1] Resilience was a negatively phrased item, scored from 1 (Agree strongly), to 5 (Disagree strongly).

2. Study 1

In study 1, we tested the one-factor hypothesis using the same sample as Huppert and So (2013).

2.1 Method

Participants. Data was acquired from the European Social Survey round 3 (2006/2007), which administered the well-being items to 43,000 individuals aged 14 and above (54% female) from 23 countries. Twenty-one cases were excluded from the analyses because they had no values for any of the items, and responses from the Hungarian sample were excluded because all values for the vitality item were missing. Thus, the overall sample was slightly reduced to 41,461 respondents. Ages ranged from 14 to 101 (M=47.35, SD=18.53), 54% were females, and 93% reported belonging to the ethnic majority. They also reported the following marital status: 51% were married, 27% were never married or in a civil partnership, 10% were widowed, 8% were divorced, 2% were in a civil partnership, 1% were separated.

Measures. Ten items were selected by Huppert and So (2013) to measure well-being (see Table 1). Positive emotion (i.e. happiness), emotional stability (i.e. calmness), vitality and resilience were scored in one direction, while the rest of the items were scored in the opposite direction. Specifically, happiness was measured on a scale from 0 (extremely unhappy) to 10 (extremely happy). Calmness and vitality were measured on a scale from 1 (none or almost none of the time) to 4 (all or almost all of the time). Competence, engagement, meaning, optimism, positive relationships and self-esteem were measured on a scale from 1 (agree strongly) to 5 (disagree strongly). The resilience item used the same agreement scale, but it was reverse-worded (see Table 1).

Table 1 Items used in Study 1 (Huppert & So, 2013) and Study 2

Construct	Items used in Study 1	Item used in study 2
Competence	Most days I feel a sense of accomplishment from what I do	Most days I feel a sense of accomplishment from what I do
Emotional stability	(In the past week) I felt calm and peaceful	In the past week, I felt calm and peaceful
Engagement	I love learning new things	I love learning new things
Meaning	I generally feel that what I do in my life is valuable and worthwhile	I generally feel that what I do in my life is valuable and worthwhile
Optimism	I am always optimistic about my future	I am optimistic about my future
Positive emotion	Taking all things together, how happy would you say you are?	Taking all things together, how happy would you say you are?
Positive relationships	There are people in my life who really care about me	There are people in my life who really care about me
Resilience	When things go wrong in my life it generally takes me a long time to get back to normal. (reverse score)	When things go wrong in my life, it generally takes me a short time to get back to normal
Self-esteem	In general, I feel very positive about myself	In general, I feel positive about myself
Vitality	(In the past week) I had a lot of energy	In the past week, I felt energetic

Analysis. Confirmatory factor analyses were carried out using the Lavaan package (version .5–18, Rosseel, 2012) in R version 3.2.0 (R Core Team, 2015). The following four models were specified: (1) the two-factor model presented in Huppert and So's (2013) study; (2) a one-factor model with no method correction; (3) a bifactor model (a) where items loaded on a general factor and two specific factors corresponding to the two factors in Huppert and So's (2013) study; (4) a bifactor model (b) where items loaded on a general factor and two specific factors each with positively and negatively scored items, respectively. If the latter bifactor model (b) fits the data better than the bifactor model (a), it would suggest that any systematic variance, not accounted for by the general factor, results from method effects rather than substantive factors. In both bifactor models all factors were uncorrelated (Holzinger & Swineford, 1937).

We used Robust Maximum Likelihood (MLR) due to some deviations from multivariate normality: Mardia's statistic of multivariate kurtosis > 5 (Bentler, 2006).

Furthermore, fewer than 7% of the cases exhibited any missing data. In total, missing values comprised approximately only 1% of the entire dataset and were estimated using full information maximum likelihood (Graham, 2009).

Several fit indices were used to evaluate how well each model fit the data. Specifically, to evaluate the fit of each model, we used the Yuan-Bentler chi-square, the Comparative Fit Index (CFI), the Standardized Root Mean Square Residual (SRMR), and the Root Mean Square Error of Approximation (RMSEA). CFI values above .90 indicate acceptable fit, while values above .95 indicate good fit; SRMR values below .06 indicate good fit; RMSEA values below .08 indicate acceptable fit and values below .06 indicate good fit (Brown, 2015). To compare the fit of competing models we used the Bayesian Information Criterion (BIC), which produces lower values for better fitting models.

Additionally, the average variance explained by each factor was compared to the variance shared by the factors to assess discriminant validity. The average variance explained by each factor is equivalent to the average of its factor loadings squared, and indicates the percentage of error-free variance in each factor. Issues of discriminant validity are identified if the variance shared by two factors is greater than the variance explained by each factor (Fornell & Larker, 1981; Hair et al., 2010). Finally, McDonald's omega hierarchical (ωh) was used to estimate the reliability of the factors, as this was found to perform better than Cronbach's α (Zinbarg, Revelle, Yovel & Li, 2005).

2.2 Results

All items were recoded in the same direction prior to the analyses. Fit indices are presented in Table 2. The two-factor model showed acceptable fit. However, the two factors were highly correlated (r = .76) and the average variance explained by each factor (30% and 29%) was lower than the variance shared between the two factors (57%), indicating poor discriminant validity.

The one-factor model without any method correction showed poor fit, as evidenced by a CFI of .867. On the other hand, the bifactor models fit the data well. In both bifactor models, items generally showed comparatively higher loadings on the general factor than the specific factors. Furthermore, while the general factor exhibited adequate reliability, the specific factors exhibited extremely low reliability (see Tables 3 and 4). The bifactor model (b) fitted the data better than model (a), as evidenced by a lower BIC, thus supporting the hypothesis that the systematic variance not explained by the general factor is produced by method effects. Finally, despite having recoded all items in the same direction, some negative loadings were found on the specific factors, but these were generally very low, in the .20 range.

Table 2. Fit indices (Study 1)

Model	YB X^2	df	CFI	SRMR	RMSEA (90% CI)	BIC
One factor	7097.72*	35	.867	.044	.070(.069-.071)	1,031,773
Two factors	4733.17*	34	.912	.034	.058(.056-.059)	1,028,742
Bifactor (a)	2143.31*	25	.960	.023	.045(.044-.047)	1,025,473
Bifactor (b)	1720.34*	25	.968	.020	.040(.039-.042)	1,025,004

Note: YB X^2 = Yuan–Bentler scaled chi square, df = degrees of freedom, CFI = Comparative Fit Index, SRMR = Standardized Root Mean Square Residual, RMSEA = Root Mean Square Error of Approximation, CI = Confidence Intervals, BIC = Bayesian Information Criterion.
* $p < .001$

Table 3. Factor loadings and reliability for the bifactor model (a)

Item content	General	One	Two
1. Happiness	0.53		0.11
2. Vitality	0.54		0.30
3. Calmness	0.50		0.33
4. Resilience	0.40		0.13
5. Optimism (R)	0.70		-0.22
6. Self-esteem (R)	0.64		-0.26
7. Engagement (R)	0.31	0.36	
8. Competence (R)	0.48	0.43	
9. Meaning (R)	0.45	0.40	
10. Relationships (R)	0.29	0.23	
Reliability (McDonald's ωh)	.72	.29	.01

Note: all items were coded as positive; the specific factors reflect the factors in Huppert and So's (2013) study.

Table 4. Factor loadings and reliability for the bifactor model (b)

Item content	General	One	Two
1. Happiness	0.49		0.22
2. Vitality	0.44		0.39
3. Calmness	0.37		0.55
4. Resilience	0.34		0.23
5. Optimism (R)	0.72	0.40	
6. Self-esteem (R)	0.61	0.19	
7. Engagement (R)	0.41	-0.20	
8. Competence (R)	0.59	-0.25	
9. Meaning (R)	0.56	-0.25	
10. Relationships (R)	0.35	-0.17	
Reliability (McDonald's ωh)	.71	.00	.20

Note: all items were coded as positive; the specific factors reflect positively and negatively scored items.

2.3 Discussion

The bifactor model with positive and negative specific factors fit the data best. Furthermore, in both bifactor models, only the general factor was reliable, while specific factors exhibited extremely low reliabilities. In most cases, items exhibited higher loadings on the general factor. The only exceptions were the engagement item in the bifactor model (a) and the calmness item in the bifactor model (b). However, these exceptions were not consistent across the two bifactor model specifications, and were therefore not expected to generalize across studies. Taken together these results indicate that the specific factors may result from different item wording, rather than substantive factors. In other words, the results corroborated previous findings showing that different well-being indicators have a strong general factor.

3. Study 2

To further test whether the two-factor model was a result of method variance, we conducted a second study in a new sample and used the same rating scale across all items. In the absence of method variance caused by different rating scales, we expected one-factor model to fit the data well, without any method correction.

3.1 Method

Participants. Data were collected from the Amazon Mechanical Turk panel (see Buhrmester, Kwang & Gosling, 2011, for an evaluation of Mechanical Turk participants). In total, 573 responses were recorded from U.S. residents. Since few responses were incomplete, the final sample included 560 participants. Only people whose previous work had been approved more than 95% of the time were able to access the questionnaire (Peer, Vosgerau & Acquisti, 2014). The 10-item scale was included in a larger 15-minute survey containing other well-being items. A small monetary incentive was used to encourage participation (approx. $0.50). A screening question was included toward the end of the survey to assess the participants' attention: "I read instructions carefully. To show that you are reading these instructions, please leave this question blank" (Maniaci & Rogge, 2014). Around 8% of participants failed this screening question and were therefore omitted from further analyses. Additionally, seven participants were excluded because their response patterns to the larger questionnaire showed no variation, as evidenced by an inter-item standard deviation of 0. Thus, the final sample comprised 507 participants. Ages ranged from 19 to 77 years (M = 38.01, SD = 13.31), 63% were female, 79% were white, 8% were black, 4% were Asian or Pacific Islander, 4% were mixed, 3% were Hispanic, 1% were Native American. With regards to marital status, 38% were married, 34% were single (never married), 14% were living with a partner, 10% were divorced, 2% were widowed and 1% were separated. With regards to educational attainment, 42% had an undergraduate degree, 36% had some college credit, 10% were high-school graduates, 7% had a masters degree, 1% had a doctorate degree, 1% had a professional degree (e.g. MD), 1% had only some high-school experience.

Measure. The scale from study 1 was revised. All items were scored in the same direction and used the same rating scale from 1 (not at all) to 5 (extremely). Few items were slightly adapted to fit this rating scale (see Table 1). For example, the terms "always" and "very" were omitted from the optimism and self-esteem items, respectively, as they would be confusing when combined with the rating scale.

Similarly, to measure vitality, the item “I had a lot of energy” was changed to “I feel energetic.” Both terms, having energy and feeling energetic, were expected to adequately tap into vitality, because they were both included in the previously-validated vitality scale (Ryan & Frederick, 1997). Finally, to avoid order effects, the items were presented in a random order.

Analysis. Consistent with study 1, the same statistical package, model estimation and fit indices were used. Since all items were measured with the same rating scale, a single unidimensional model was specified, and no method correction was required.

3.2 Results

The two-factor model and the one-factor model fit the data equally well (see Table 4). The BIC slightly favored the one-factor model, although the difference between the two models was extremely small. In other words, the two models appeared to fit the data equally well. However, the two-factor model once again exhibited poor discriminant validity: the two factors were very highly correlated (r = .97) and their shared variance (94%) exceeded the average variance explained by each individual factor (46% and 61%). Loadings for the one-factor model ranged from .38 to .88, and the factor explained on average 54% of the variance in the items.

Table 4. Fit indices (Study 2)

Model	YB X^2	df	CFI	SRMR	RMSEA (90% CI)	BIC
One factor	100.14*	35	.972	.030	.061(.048-.073)	12,537
Two factors	96.05*	34	.973	.030	.060(.048-.073)	12,538

Note: YB X^2 = Yuan–Bentler scaled chi square, df = degrees of freedom, CFI = Comparative Fit Index, SRMR = Standardized Root Mean Square Residual, RMSEA = Root Mean Square Error of Approximation, CI = Confidence Intervals, BIC = Bayesian Information Criterion.
* $p < .001$

4. General discussion

After accounting for method effects, well-being indicators included in both studies exhibited a strong unidimensional structure. While a two-factor model also showed adequate fit to the data, it exhibited issues of discriminant validity, as the variance shared by the two factors was higher than the variance explained by each individual factor. Taken together, these results suggest that the components of well-being included in Huppert and So's (2013) scale may be adequately explained by a single factor. Based on these findings, if each construct was measured with a multi-item scale, rather than a single-item scale, we would expect a single higher-order factor model to fit the data well. Indeed, the items used in both studies were chosen to measure 10 different constructs, and only one item per construct was available. If the measure had comprised several items per construct, we would expect that measure to be multidimensional at the lower-order level (e.g. with 10 lower-order factors), but with a single general or higher-order factor.

Furthermore, the results suggest that a single scale score would be justified for the items in the current well-being scale. If composite scale scores are to be used in future studies, we recommend using the revised scale in study 2, because this version requires no adjustments for method effects.

Clearly, a single-factor higher-order model is simpler and more parsimonious than a two- or three-factor higher-order model. Often, more parsimonious models (e.g. with one factor) tend to show worse fit than more complex models (e.g. with several factors). As we estimate more parameters, thus increasing complexity, model fit tends to improve (e.g. Brown, 2015). Nevertheless, as long as a simple model fits the data adequately, it may be preferable over more complex models (e.g. Brown, 2015; Kim & Mueller, 1978). Indeed, parsimony is an aspect that is highly valued in science (Kline, 2000). Therefore, our preference of a unidimensional model is not merely based on improvements in fit but on conceptual parsimony.

The current results are comparable to previous findings in the literature. They are consistent with findings showing that a single higher-order factor adequately explains the relationships among positive feeling and positive functioning indicators (e.g. Disabato et al., in press; Gallagher et al., 2009). However, some inconsistencies can also be identified in the literature. Some studies supported a strong general factor in a bifactor model, but a general factor, on its own, did not fit the data adequately (e.g. Jovanovic, 2015). Nevertheless, taken together, the findings are consistent in showing that indicators of well-being seem to reflect a single general dimension, more than a distinction between positive feeling and positive functioning (e.g. Chen et al., 2013; de Bruin & Du Plessis, 2015; Gallagher et al., 2009; Jovanovic, 2015).

Based on current findings, we outline the following limitations and directions for future research. It must be emphasized that current results relate to the factor structure of the well-being indicators in one specific scale, rather than the structure of well-being in general. Furthermore, a specific 5-point rating scale was used in study 2, and the choice of rating scale may have influenced the results. Therefore, it would be useful to replicate the current studies using a variety of well-being measures, using a variety of rating scale formats. Furthermore, since this and previous studies have focused on Western samples, a range of samples from a variety of cultures should be included in future studies, as the structure of well-being may change in different cultures (e.g. Russell, 1980). Finally, the relations between the well-being scale and other measures should be explored. Specifically, it would be useful to examine whether separate higher-order factors are differentially related to external measures. If similar relations with external variables were found (e.g. Disabato et al., in press), this would further support the use of single higher-order or general factor.

References

Alessandri, G., Vecchione, M., Eisenberg, N., & Łaguna, M. (2015). On the Factor Structure of the Rosenberg (1965) General Self-Esteem Scale. *Psychological Assessment*, *27*, 621-635.

Bachman, J. G., & O'Malley, P. M. (1986). Self-concepts, self-esteem, and educational experiences: The frog pond revisited (again). *Journal of Personality and Social Psychology, 50*(1), 35-46.

Bentler, P. M. (2006). *EQS 6 Structural Equations Program Manual*. Encino, CA: Multivariate Software, Inc.

Brown, T.A. (2015). *Confirmatory factor analysis for applied research*. London: The Guilford Press.

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science, 6*(1), 3–5.

Chen, F. F., Jing, Y., Hayes, A., & Lee, J. M. (2013). Two concepts or two approaches? A bifactor analysis of psychological and subjective well-being. *Journal of Happiness Studies, 14*(3), 1033-1068.

de Bruin, G. P., & du Plessis, G. A. (2015). Bifactor analysis of the mental health continuum–short form (MHC–SF). *Psychological Reports, 116*(2), 438-446.

Diener, E. (1984). Subjective well-being. *Psychological Bulletin, 95*, 542–575.

Disabato, D.J., Goodman F.R., Kashdan T.B., Short J.L., & Jarden A. (in press). Different types of well-being? A cross-cultural examination of hedonic and eudaimonic well-being. *Psychological Assessment*. doi: 10.1037/pas0000209

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*, 39-50.

Gallagher, M. W., Lopez, S. J., & Preacher, K. J. (2009). The hierarchical structure of well-being. *Journal of Personality, 77*(4), 1025-1050.

Graham, J. W. (2009). Missing data analysis: Making it work in the real world. *Annual Review of Psychology*, *60*, 549-576.

Hair, J.F., Black, W.C., Babin, B.J, Anderson, R.E. (2010). *Multivariate data analysis*: *A global perspective*. New York: Pearson.

Huppert, F. A., & So, T. T. (2013). Flourishing across Europe: Application of a new conceptual framework for defining well-being. *Social Indicators Research*, *110*(3), 837-861.

Jovanović, V. (2015). Structural validity of the Mental Health Continuum-Short Form: The bifactor model of emotional, social and psychological well-being. *Personality and Individual Differences, 75*, 154-159.

Holzinger, K. J., & Swineford, F. (1937). The bi-factor method. *Psychometrika, 2*, 41-54.

Keyes, C. L. (2002). The mental health continuum: From languishing to flourishing in life. *Journal of Health and Social Behavior*, *43*(2), 207-222.

Keyes, C. L., Shmotkin, D., & Ryff, C. D. (2002). Optimizing well-being: the empirical encounter of two traditions. *Journal of Personality and Social Psychology*, *82*(6), 1007-1022.

Kim, J. O., & Mueller, C. W. (1978). *Factor analysis: Statistical methods and practical issues*. London: Sage.

Kline, P. (2000). *The handbook of psychological testing*. New York: Routledge.

Linley, P. A., Maltby, J., Wood, A. M., Osborne, G., & Hurling, R. (2009). Measuring happiness: The higher order factor structure of subjective and psychological well-being measures. *Personality and Individual Differences*, *47*(8), 878-884.

Maniaci, M. R., & Rogge, R. D. (2014). Caring about carelessness: Participant inattention and its effects on research. *Journal of Research in Personality, 48*, 61-83.

Marsh, H. W. (1996). Positive and negative global self-esteem: A substantively meaningful distinction or artifactors? *Journal of Personality and Social Psychology, 70*(4), 810-819.

Peer, E., Vosgerau, J., & Acquisti, A. (2014). Reputation as a sufficient condition for data quality on Amazon Mechanical Turk. *Behavior Research Methods, 46*(4), 1023-1031.

R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria. URL http://www.R-project.org/.

Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software, 48*(2), 1-36.

Russell, J.A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39*(6), 1161-1178.

Ryan, R. M., & Deci, E. L. (2001). On happiness and human potentials: A review of research on hedonic and eudaimonic well-being. *Annual Review of Psychology, 52*(1), 141-166.

Ryan, R. M., & Frederick, C. (1997). On energy, personality, and health: Subjective vitality as a dynamic reflection of well-being. *Journal of Personality, 65*(3), 529-565.

Ryff, C. D. (1989). Happiness is everything, or is it? Explorations on the meaning of psychological well-being. *Journal of Personality and Social Psychology*, *57*(6), 1069-1081.

Schmitt, N., & Stults, D. M. (1985). Factors defined by negatively keyed items: The result of careless respondents? *Applied Psychological Measurement, 9*(4), 367-373.

Woods, C. M. (2006). Careless responding to reverse-worded items: Implications for confirmatory factor analysis. *Journal of Psychopathology and Behavioral Assessment, 28*(3), 186-191.

Zinbarg, R. E., Revelle, W., Yovel, I., & Li, W. (2005). Cronbach's α, Revelle's β, and McDonald's ωh: Their relations with each other and two alternative conceptualizations of reliability. *Psychometrika, 70*, 123-133.