The German Version of the PERMA-Profiler: Evidence for Construct and Convergent Validity of the PERMA Theory of Well-Being in German Speaking Countries

Martin Wammerl^{1,2} · Johannes Jaunig^{2,3} · Thomas Mairunteregger^{1,2} · Philip Streit²

Published online: 13 November 2019

© The Author(s) 2019

Abstract

With the PERMA theory, Seligman (2011) postulates that well-being consists of five independently measurable factors: Positive Emotions (P), Engagement (E), Positive Relationships (R), Meaning (M) and Accomplishment (A). The PERMA-Profiler provides the first questionnaire, which measures all five well-being domains in an economical and reliable way. In order to test the validity of the questionnaire in German speaking countries, a German version of the PERMA-Profiler was developed and evaluated in a large sample ($N=854$). The results provide evidence for acceptable reliability, very good construct validity (factorial and convergent) and first indications for measurement invariance, for both gender and nationality. Compared to three theoretically competing models, the inter-correlated Five-Factor Model turned out to be the most appropriate statistical model to describe the collected data. It revealed the best trade-off between model fit, parsimony and theoretical interpretability. Our results support the hypothesis of a multidimensional PERMA theory, which gives a closer insight in at least some of the building blocks of well-being. Therefore, the PERMA theory can be seen as a useful extension to a unidimensional subjective well-being approach. Like the English original, the German version of the PERMA-Profiler allows to measure well-being economically across multiple well-being domains. Therefore, the PERMA-Profiler can be recommended as a valid well-being screening instrument for the German speaking adult population.

Keywords PERMA · PERMA-profiler · Well-being · Confirmatory factor analysis · Measurement invariance · Bi-factor · German language · Questionnaire

✉ Martin Wammerl
martin.wammerl@edu.uni-graz.at

¹ University of Graz: Institute of Sociology, Universitätsstraße 15, 8010 Graz, Austria

² Institute of Positive Psychology and Mental Coaching, Walter-Goldschmidt-Gasse 25, 8042 Graz, Austria

³ University of Graz: Institute of Sports Science, Mozartgasse 14, 8010 Graz, Austria

The research on the measurement of subjective well-being¹ goes back to the beginning of the twentieth century. At this time, the first self-report scales were developed in order to assess different facets and dimensions of subjective well-being (SWB). In his comprehensive review of the literature on SWB, Wilson (1967) concluded that socio-economic factors (i.e. age, health, education, socio-economic status, marital status) as well as different psychological variables (i.e. extraversion, optimism, religiousness, self-esteem and intelligence) influence the subjective happiness ratings. Since that time the research on subjective well-being aimed to focus less on demographic characteristics and turned towards the understanding of the processes that underlie SWB. One important component of SWB was identified through the research of Bradburn (1969). His research on coping strategies when facing difficulties in daily life showed that happiness was one of the most important variables involved. To experience happiness, the positive emotions must prevail. Furthermore, the studies have shown that there is a distinction between positive and negative emotions and that these are to be considered as separate variables. Later, social scientists have focused on what leads people to evaluate their lives in positive terms. In this context, life-satisfaction, which relies on the cognitive judgement of what is the good life, was added as another component of subjective well-being (Shin and Johnson 1978; for an overview: Diener 1984). Diener and Suh (1997) concluded that “both affect and reported satisfaction judgments represent people’s evaluations of their lives and circumstances” (p. 200).

Today, well-being has entered the public discourse with scientists and public figures alike demanding better and reliable measures of people’s well-being (Layard 2010). Based on Greek philosophy, the two main theoretical approaches discussed with regard to well-being and its measurement are eudaimonia and hedonism (Delle Fave et al. 2011; Dodge et al. 2012; Stoll 2014). “The hedonic view equates happiness with pleasure, comfort, and enjoyment [while] the eudaimonic view equates happiness with the human ability to pursue complex goals which are meaningful to the individual and society” (Delle Fave et al. 2011, p. 5). Depending on the underlying theories a large number of questionnaires were developed in order to measure well-being or specific well-being components. One of the most prominent scales, mostly based on the hedonistic approach, is the Satisfaction with Life Scale (Diener et al. 1985), which gives a global estimation of a person’s satisfaction with life. A multifactorial approach to measure well-being, which relies on an eudaimonic background is the theory of psychological well-being (PWB) and the corresponding questionnaire, the Psychological Well-Being-Scale (PWBS), of Ryff and Keyes (1995) which postulates that well-being consists of six fundamental factors (Autonomy, Environmental Mastery, Personal Growth, Positive Relations with Others, Purpose in Life and Self-Acceptance). Even though the PWBS is still frequently used, the recent critique of the validity of the postulated factor structure (Abbott et al. 2006) forced research to make a determined effort to develop new, valid, reliable and economic well-being questionnaires to measure multiple dimensions of PWB. The Flourishing Scale (Diener et al. 2009) subsumes different PWB theories and provides a single score, measuring self-perceived success in areas such as relationships, self-esteem, purpose and optimism. Another important questionnaire is the so called Inventory of Thriving (Su et al. 2014), which aims to measure a broad range of PWB components and to predict important health outcomes. The third important approach to measure different dimensions of PWB is the PERMA theory (Seligman 2011).

¹ For detailed information on the historical background of the well-being research and the difficulties related to defining the concept of well-being, see Dodge et al. (2012).

When Seligman (2002) introduced his theory of Authentic Happiness, he posited well-being to be consisting of three major elements: The first is *Positive Emotions* or the pleasant life, which is all about feelings such as pleasure, joy, gratitude or hope. The second is *Engagement* or flow and refers to a psychological state in which a person becomes one with a task and more or less loses the sense for time and self-consciousness. Finally, *Meaning* refers to a sense of purpose in life or serving something higher than oneself. Seligman (2002) theorized, if people scored high on those three dimensions, their life satisfaction should be high as well or that they would experience happiness. With this approach, Seligman manages to incorporate both, a hedonic and eudaimonic view of well-being into his theory (Henderson and Knight 2012). Later, in his book *Flourish*, Seligman (2011) reports certain limitations of the self-report assessment of well-being, which is influenced by the current mood while filling out the questionnaire (Huppert et al. 2009; Seligman 2011). Additionally, authentic happiness stands and falls with life-satisfaction measures and “comes dangerously close to Aristotle’s monism because happiness is operationalized, or defined, by life satisfaction” (Seligman 2011, p. 16) alone. Also, the three mentioned elements, which constitute authentic happiness, do not incorporate all there is that “people choose for their own sake” (p. 14). To circumvent those issues, Seligman replaced life-satisfaction with well-being as the goal for positive psychology and adds *Positive Relationships* and *Accomplishment* to the three elements already in place, which now together constitute the PERMA model of well-being or flourishing. Seligman argues that the advantage of using well-being over life-satisfaction is the fact that well-being, as Seligman defines it, is a construct of five dimensions, which can “each [be] contributing to well-being, but not defining well-being” (p. 15), while life-satisfaction has to stand on its own.

The merit of Seligman’s PERMA model is still in debate. Kashdan (2017) and Goodman et al. (2017) argued that PERMA is redundant because of its strong correlation with subjective well-being and thereby not yielding a new kind of well-being or giving any new insights, respectively. Seligman (2018) replied that his claim did not constitute PERMA as a new kind of well-being but rather gives a closer insight in at least some of the building blocks of SWB. Therefore, the PERMA theory can be seen as a useful extension of SWB, especially when thinking of the development of specific interventions, which should selectively address one of these building blocks in the context of psychotherapy, coaching or counseling.

1 Operationalizing PERMA

There has been some additional critique (Krueger 2012; Van Zyl 2013) regarding the lack of empirical evidence for the PERMA model as it was presented in Seligman’s (2011) book *Flourish*. However, by now there has been an approach made to shed empirical light on this issue. This first approach is by Butler and Kern (2016), who created a brief instrument that measures all five PERMA domains: The PERMA-Profil. It consists of 15 items (three per factor) representing the five PERMA factors: Positive Emotions, Engagement, Positive Relationships, Meaning and Accomplishment.

In the studies of Butler and Kern (2016), the data supported the five-factor structure of the PERMA model. Confirmatory factor analysis revealed that the inter-correlated Five-Factor Model adequately fit the data and that the five factors were generally reliable. The PERMA theory (but not the PERMA-Profil) was tested within a Higher-Order structure, which showed superior model fit compared to a Single-Factor solution, but was not contrasted to other model types (Coffey et al. 2014). As mentioned above, recent psychometric analysis of

the PERMA theory showed that SWB and PERMA share over 88% of their variance (Goodman et al. 2017). The study of Goodman et al. has some substantial limitations, where all questionnaires were scored on the same 5-point Likert scale and the CFAs were conducted at facet level rather than on item level. These limitations could bias towards the proposed very high latent correlation between SWB and PERMA. Chen et al. (2013) found for the PWBS, structured as a Bi-Factor Model, that “the specific factors demonstrated incremental predictive power, independent of the general well-being factor. These results suggest that psychological well-being and subjective well-being are strongly related at the general construct level, but their individual components are distinct” (p. 1033). Bi-Factor Models are more and more on the rise in personality and particularly in well-being research (Reise 2012). There is growing evidence that well-being is represented best in a Bi-Factor Model. In recent years the Mental Health Continuum – Short Form (MHC-SF; Keyes et al. 2008) was extensively evaluated using competing models and revealed a consent for a Bi-Factor Model structure (e.g. De Bruin and Du Plessis 2015; Hides et al. 2016; Jovanović 2015; Lamborn et al. 2018; Schutte and Wissing 2017; Silverman et al. 2018). This provides evidence for a general well-being factor, which could explain the inter-correlations of the PERMA factors. However, the divergence between the results of Butler and Kern, Coffey et al., and Goodman et al. leads to the open question which factorial model represents the PERMA theory of well-being best.

2 Present Study

With this study, we try to expand on Butler and Kern (2016) initial findings and want to examine the validity of the PERMA theory in German speaking countries by developing and testing a German version of the PERMA-Profiler. To our knowledge, there has not been a study, with the attempt to adapt the PERMA-Profiler with regard to content and language for a German speaking population. In order to test the construct validity of the measure, we included two measures which are widely used in German speaking countries: The 18-item version of the PWBS (Ryff and Keyes 1995) and the Depression Anxiety and Stress Scale (Lovibond and Lovibond 1995). Both measures have not been used previously to assess the PERMA-Profilers validity. In order to test the factorial validity and the latent variable model of the PERMA theory, we aim to test the PERMA-Profiler questionnaire with a confirmatory factor analysis approach. In the present study, we contrast four different model alternatives like in previous studies in well-being research (e.g. De Bruin and Du Plessis 2015 for the MHC-SF): The original inter-correlated Five-Factor Model of Butler and Kern (2016), a Single-Factor Model, a Higher-Order Model and a Bi-Factor Model. The Single-Factor Model excludes the PERMA factors and states that all items load on one single well-being factor. The Higher-Order Model adds a second-order well-being factor, whose effect is fully mediated by the five PERMA-factors. This type of model pictures a “bottom-up” approach of prioritizing the PERMA factors over the well-being factor (Beaujean 2015). The Bi-Factor Model, which can also represent the existence of both a general well-being factor and the five PERMA factors, could be a plausible alternative to the Higher-Order Model. In contrast to the Higher-Order Model, the Bi-Factor Model reflects the assumption that the associations of the general well-being factor with the observed item responses is direct and independent of the associations of the PERMA factors with the item-responses. In this context the PERMA factors are deducted from the responses on the questionnaire, but also account for additional variance in relation to a general well-being factor. Consequently, it assumes a “top-down” approach on well-being (Beaujean 2015). This

model would be in line with the recent findings for the factor structure of the MHC-SF (Keyes et al. 2008) mentioned above. The present study aims to test, which of these four models describes best the collected data in German speaking countries. Psychometric model comparisons pursue the hypothesis that the PERMA-Profiler measures a multidimensional construct. Because of the inconsistent data considering the factorial structure we have no specific hypothesis which of the multidimensional models (Five-Factor, Higher-Order or Bi-Factor) is superior.

Recent research shows that, while men and woman do not substantially differ with regard to their average well-being, woman are over-represented in the extreme ends of the well-being spectrum due to the fact that they experience positive and negative emotions more frequently and more intensely (Diener and Ryan 2009). That is why we included an investigation of the measurement invariance across gender. Additionally, we also compared the measurement invariance of nationalities because of the existing linguistic differences between Austrian and German citizens (Ammon 1995). For a distinct interpretation of the model and group comparisons on latent mean differences it is crucial to deliver empirical evidence that the underlying PERMA structure is equivalent across those groups.

3 Method

3.1 Sample

The sample consisted of 854 German speaking individuals (83.6% women) with a mean age of 28.9 years ($SD = 10.6$, range: 18–65), who completed the whole online survey. Inclusion criterion was legal age or older (≥ 18 years), citizenship (AUT, DE and CH) and German as a first language. The recruitment was based on a large internet initiative, calling for participants on various German speaking social media platforms and web pages of Austrian psychological institutes. Data were collected from August 19, 2014 to November 20, 2014. Regarding the educational level, two participants (0.2%) had no secondary education, 19 (2.2%) finished compulsory school (i.e., Pflichtschule), 54 (6.3%) had concluded an apprenticeship (i.e., Lehre, Ausbildungsberuf), 390 (45.7%) had a general qualification for university entrance (i.e., Matura, Abitur), 61 (7.1%) finished vocational school (i.e., Kolleg, Fachhochschulen), and 328 (38.4%) had a university degree (Universität, Fachhochschule). Considering the nationality of the sample, 520 participants were Austrian followed by 317 Germans and 17 participants from Switzerland.

3.2 Procedure

The survey included a sociodemographic part and three questionnaires, in particular the German version of the PERMA-Profiler, and two validation questionnaires, namely the German version of the Depression-Anxiety-Stress-Scale (DASS; Nilges and Essau 2015) and the German 18-item short-version of the PWBS (Bartkowiak 2008). The order of the single questionnaires was the same for all participants, starting with the sociodemographic questions, followed by the PERMA-Profiler, DASS and finishing with the PWBS. The average time for completing the questionnaire was 34 min. The data was collected through an open source survey application (LimeSurvey Project Team 2015). Considering the data quality of online surveys, several studies concluded

that there is no substantial difference between both online and paper surveys regarding social desirability (Dodou and de Winter 2014), as well as psychometric properties and factor structure (Wolgast et al. 2014). All participants received an informed consent form on the first page of the online survey. No identifying information about the participants was collected.

3.3 Measure

The *PERMA-Profiler* is a brief multidimensional measure of psychological well-being that allows individuals and organizations to assess and monitor well-being in terms of Seligman's (2011) PERMA theory. According to Butler and Kern (2016), the development of the PERMA-Profiler involved three parts. First, a bank of over 700 items, which were theoretically relevant to the five PERMA domains, was created. An expert rating procedure, to identify the items, which are valid to represent each domain, reduced the item pool to a final item bank of 109 questions (33 Positive Emotions, 23 Engagement, 21 Relationships, 15 Meaning and 17 Accomplishment). Furthermore, the authors chose to include positively worded items only, which lead to a further reduction to 70 items. Due to theoretically inconsistent factorial loadings of reversed-scored items, this approach avoids method-induced biases (see Dunbar et al. 2000). Finally, the sample was split into two halves randomly. Each half was analyzed with an exploratory principal component analysis, specifying a five-factor structure. For the final item set, the authors selected those that consistently appeared in the specific factor in both sub-samples (Butler and Kern 2016). The final questionnaire consists of 23 items with three items representing each of the five PERMA components. In addition to the 15 PERMA items the questionnaire includes eight filler items, which are aimed to disrupt response tendencies and providing additional information about the participants. The eight filler items comprise one item assessing overall happiness, three negative emotion items assessing sadness, anger, and anxiety, one item assessing loneliness, and three items assessing self-perceived physical health. The overall happiness question serves as a comparison with other population-based surveys. Each item is scored on an 11-point rating scale, anchored by 0 (never) to 10 (always) or 0 (not at all) to 10 (completely). The three item-scores of each domain are averaged to produce a single domain score ranging from 0 to 10 (higher scores indicate greater well-being). A total score can be calculated by summarizing the score of the 15 PERMA-items. The extensive studies by Butler and Kern demonstrated an acceptable model fit, internal and cross-time consistency, and evidence for content and convergent validity.

The German version of the *DASS* (Nilges and Essau 2015) is a set of three self-report scales designed to measure the negative emotional states of Depression, Anxiety and Stress. Each of the three DASS scales contains 14 items, divided into subscales of two to five items. The Depression subscale assesses dysphoria, hopelessness, devaluation of life, self-deprecation, lack of interest/involvement, anhedonia, and inertia. The Anxiety subscale assesses autonomic arousal, skeletal muscle effects, situational anxiety, and subjective experience of anxious affect. The Stress subscale is sensitive to levels of chronic non-specific arousal. It assesses difficulty relaxing, nervous arousal, and being easily upset/agitated, irritable/over-reactive and impatient. Participants are asked to use 4-point severity/frequency scales (0 = *Did not apply to me at all* to 3 = *Applied to me very much, or most of the time*) to rate the extent to which they have experienced each state over the past week. Scores for Depression, Anxiety and Stress are calculated by summing the scores for the relevant items. The reliability of the German version

of the DASS could be confirmed to be sufficiently high for all subscales (Depression: $\alpha = .88$, Anxiety: $\alpha = .76$ and Stress: $\alpha = .86$). The examination of the convergent validity showed high correlations of the subscales with the corresponding scales of Beck's Depression Inventory (BDI; $r = .76$) and Beck's Anxiety Inventory (BAI; $r = .68$).

The 18-item short-form of the *Psychological Well-Being-Scale* (PWBS, Ryff and Keyes 1995) is a well-established questionnaire and includes a series of statements with three items reflecting one of the six areas of psychological well-being: Autonomy, Environmental Mastery, Personal Growth, Positive Relationships, Purpose in Life and Self-Acceptance. Respondents rate statements on a scale of 1 to 6, with 1 indicating strong disagreement and 6 indicating strong agreement. The long version shows sufficient test-retest reliability ($r_{tt} = .81$ to $r_{tt} = .88$; Ryff 1989) and convergent validity. Subscale correlations are sufficiently high with negative functioning ($r = -.31$ to $r = -.73$) and with positive functioning $r = .25$ to $r = .73$). The internal consistencies of the 18-item German version (Bartkowiak 2008; $\alpha = .41$ to $\alpha = .70$) are low but similar to the values provided by Ryff and Keyes ($\alpha = .33$ to $\alpha = .56$).

3.4 Translational Process and Cross-Cultural Adaption

The translation of the PERMA-Profiler was based on the international guidelines for cross-cultural adaption of self-report measures (Beaton et al. 2000). The translation process consisted of following five stages: Initial translation (Stage 1), synthesis of the translations (Stage 2), back translation (Stage 3), expert validation (Stage 4), and test of the pre-final version (Stage 5). In the first stage, two forward translations from English into German were developed by two bilingual translators, respectively, who are native in German. One translator was aware of the PERMA theory (psychologist), the other translator had no prior knowledge about the questionnaire. The different qualifications of the translators should lead to an accurate detection of ambiguous meanings in the original questionnaire. In Stage 2 both translators and a recording observer synthesized the two questionnaire versions into one preliminary version of the German translation. The preliminary version was subject to a back-translation (Stage 3), which was done by two independent English native speakers. Both back-translators were not informed about the PERMA theory. Out of the five resulting questionnaire versions (two German, one common German, and two back translations), the study authors, together with the translators, created a pre-final version of the questionnaire (Stage 4). The goal in this stage was to achieve semantic equivalence between the English original and the German target items. The final stage (Stage 5) consisted of a "think-aloud" cognitive interview (Ericsson and Simon 1980) with a sample of 10 adult participants. The participants were instructed to "think aloud" as they answer the questions of the pre-final version. Through this technique potential problems in understanding the intended meaning of the individual items could be detected. Based on the results of the cognitive interview, the final version for the validation procedure was generated.

3.5 Data Analyses

Confirmatory factor analyses (CFA) were used to evaluate the proposed factor structure of the 15-item PERMA model on our German speaking sample. As estimator we chose maximum likelihood estimation with robust (Huber-White) standard errors (MLR) to address the potential problem of result bias through non-normal distributed data (Yuan and Bentler 2000; Finney and DiStefano 2013). Reliability (internal consistency) of all three questionnaires was assessed

with Cronbach's alpha coefficient (α) and Guttman's lambda 6 (λ_6). We used Lavaan R package (version 0.6–2; Rosseel 2012) in RStudio (version 1.1.456; RStudio 2016) for CFA and measurement invariance computations, all other statistical analyses were conducted in SPSS 21. To test for normality of the data an index for acceptable limits of skewness and kurtosis of ± 2 was used (Field 2009). Four models were defined for the model comparison analysis: (a) The Single-Factor Model, in which all items load directly on a single well-being factor. (b) The Five-Factor Model, which includes the five specific PERMA factors but no general well-being factor. In this model every PERMA factor is defined through the three specific items. The five PERMA factors are inter-correlated but separate constructs. (c) The Higher-Order Model adds a general well-being factor to the factor structure. In this model the item responses fully mediate well-being through the five PERMA factors. (d) Finally, the Bi-Factor Model assumes that the item responses directly influence a general-well-being factor. The PERMA factors, which are composed of the specific items, are modelled to be independent from the general well-being factor. The four competing models are pictured in Fig. 1. To test the global goodness-of-fit of the models, the following fit indices, with corresponding cut-off values suggested by Beauducel and Wittmann (2005) and Hu and Bentler (1998, 1999) were used: Chi-Square test, as well as four goodness-of-fit indices: (a) The root mean square error of approximation (RMSEA; $\leq .06$), (b) the standardized root mean square residual (SRMR; $\leq .08$), and (c) the comparative fit index (CFI; $\geq .95$) and (d) the Tucker-Lewis index (TLI; $\geq .95$). Research on the validity of the fit indices showed that the statistical power of the Chi-Square test is highly dependent on the sample size, where “in large samples virtually any model tends to be rejected” (Bentler and Bonett 1980, p. 588). In recent simulation studies with a large sample size ($N > 800$) and similar model specifications as used in our study, the Chi-Square test reached nearly 100% power, and rejected nearly every model (Meade 2008). For this reason, we focused the evaluation of the global model fit on the four goodness-of-fit indices. For model comparisons of the competing models, we used the goodness-of-fit indices and the Bayes information criterion (BIC, Raftery 1995) and for nested models (i.e. the

Fig. 1 Measurement and structural model of the four competing PERMA models. **a** Single-Factor Model. **b** Higher-Order Model. **c** Correlated Five-Factor Model. **d** Bi-Factor Model

Higher-Order and the Bi-Factor Model) additionally the scaled Chi-Square difference test (Satorra and Bentler 2001).

In order to test for measurement invariance, gender and nationality were taken into account. Our calculations are based on the well-established procedure for multi-group CFA's with increasingly restricted nested models (Kline 2015; Vandenberg and Lance 2000). We tested for following restrictions: configural (M1; no restrictions), weak or metric (M2; adding equal factor loadings), strong or scalar (M3; adding equal intercepts), and strict or full uniqueness (M4; adding equal residual variances) measurement invariance. For model comparisons, the scaled Chi-Square difference test (Satorra and Bentler 2001) and the changes in the CFI (ΔCFI) as well as in the RMSEA ($\Delta RMSEA$) were taken into account. However, as mentioned above, the Chi-square difference test is affected by the same issues as the Chi-Square test for global model fit (Chen 2007; Little 2013). Consequently, only the following criteria were used to detect a practical lack of invariance: $\Delta CFI \geq -.01$ (Cheung and Rensvold 2002) and $\Delta RMSEA \geq .015$ (Chen 2007). The pairwise comparisons were made with its less restrictive predecessor (i.e., M1 vs. M2, M2 vs. M3, etc.).

4 Results

4.1 Confirmatory Factor Analysis

4.1.1 Global Model Fit

First, we defined the inter-correlated Five-Factor Model, derived from the studies of Butler and Kern (2016), as the Five-Factor Model. Secondly, we tested theoretically competing models and compared them with the Five-Factor Model according to the goodness-of-fit indices and the BIC. We defined the following three competing models: A Single-Factor Model, a Higher-Order Model and a Bi-Factor Model. Figure 1 illustrates the structural difference between those four models (Appendix Table 5).

Table 1 shows the model fit for the Single-Factor Model, the Higher-Order Model, the Five-Factor Model, and the Bi-Factor Model. All models showed a significant exact model test (Chi-Square test). Because of the large sample size and (minor) deviations from normality, we used the goodness-of-fit indices and the BIC as reference for comparison. Except for the Single-Factor structure, all models met the cut-off criteria and showed a good global model fit. This first result supports a multidimensional factor solution. Relating to the remaining models,

Table 1 Fit indices of the competing models

	χ^2_{MLR}	df	Scaling	RMSEA (90% CI)	SRMR	CFI	TLI	BIC
Single-factor	974.85**	90	1.402	.107 (.102–.112)	.064	.825	.796	49,818.6
Higher-order	286.80	85	1.367	.053 (.047–.059)	.042	.960	.951	48,877.9
Five-factor	220.05	80	1.371	.045 (.039–.051)	.034	.972	.964	48,821.2
Bi-factor	195.77	75	1.354	.043 (.037–.050)	.031	.976	.968	48,818.3

** $p < .01$; $N = 854$. χ^2_{MLR} - Scaled Chi-Square using maximum likelihood estimation with robust standard errors (MLR); *df*, Degrees of freedom; *Scaling*, Scaling correction factor for the Chi-Square from the MLR estimator; *RMSEA*, Root Mean Square Error of Approximation; *CI*, Confidence Interval; *SRMR*, Standardized Root Mean Square Residual; *CFI*, Comparative Fit Index; *TLI*, Tucker Lewis Index; *BIC*, Bayes Information Criterion

the Higher-Order Model showed the worst goodness-of-fit values. The Bi-Factor Model showed a better model fit than the Higher-Order Model and a slightly better model fit and BIC than the Five-Factor Model. The scaled Chi-Square difference test (Satorra and Bentler 2001) showed a significantly better model fit for the Bi-Factor Model, compared to the Higher-Order Model ($\Delta\chi^2_{10} = 86.71$, $p < .01$). Because of the non-nested structure, no Chi-Square difference test for the Five-Factor Model was possible (Steiger et al. 1985).

4.1.2 Local Model Fit

All multidimensional models (Higher-Order, Five-Factor, and Bi-Factor) showed significant variances and factor loadings ($p < .01$). Both, the factor loadings and the inter-factor correlations were invariably high for the Five-Factor Model. The Bi-Factor Model showed lower factor loadings for the specific PERMA domains. Seven out of 15 items had a factor loading $< .40$ (for details see Appendix Table 6). This difference can be explained through the assignment of two factors instead of one factor to each item. The Higher-Order Model factor loadings were comparably high as in the Five-Factor Model (Appendix Table 7).

To sum up, the Bi-Factor showed a better global model fit than the Higher-Order Model and a slightly better fit than the Five-Factor Model. The high latent inter-correlations between the PERMA factors in the Five-Factor Model as well as the high correlations between the overall happiness item and the PERMA factors might suggest an underlying general well-being factor. However, the Single-Factor Model was rejected, because it did not meet the cut off criteria for global model fit. The Bi-Factor Model and the Five-Factor Model showed nearly equally high fit indices. Because of the theoretical compatibility with the original assumptions of the PERMA theory, we chose the Five-Factor Model to represent the collected data best.

Since the Five-Factor Model met the cut off criteria for all goodness-of-fit indices, we used the unmodified model in all further calculations.

4.1.3 Measurement Invariance

In order to test the Five-Factor Model for measurement invariance, relating to gender and nationality, we focused on the changes in the CFI and the RMSEA at the pairwise comparisons on the increasingly restricted nested models: first, for male and female and second, for

Table 2 Fit indices of the Five-Factor Model separated by gender and nationality

	<i>N</i>	χ^2_{MLR}	<i>df</i>	<i>Scaling</i>	RMSEA (90% CI)	SRMR	CFI	TLI	BIC
Gender									
Female	714	187.33**	80	1.378	.043 (.036–.050)	.035	.973	.966	40,773.1
Male	140	148.02**	80	1.165	.078 (.060–.096)	.051	.943	.931	8195.2
Nationality									
Austrian	520	163.54**	80	1.358	.045 (.036–.053)	.033	.973	.964	29,879.0
German	317	137.80**	80	1.316	.048 (.036–.059)	.047	.971	.961	18,265.1

** $p < .01$; χ^2_{MLR} - Scaled Chi-Square using maximum likelihood estimation with robust standard errors (MLR); *df*, Degrees of freedom; *Scaling*, Scaling correction factor for the Chi-Square from the MLR estimator; *RMSEA*, Root Mean Square Error of Approximation; *CI*, Confidence Interval; *SRMR*, Standardized Root Mean Square Residual; *CFI*, Comparative Fit Index; *TLI*, Tucker Lewis Index; *BIC*, Bayes Information Criterion

Austrian and German participants, respectively. The Swiss participants were excluded from the analysis because of the poor attendance ($n = 17$). As illustrated in Table 2, all four subgroups showed a significant exact model test (Chi-Square test). Considering the goodness-of-fit indices, the Five-Factor Model met the cut-off-criteria for all sub-groups. The sole exception was the RSMEA, which was slightly higher in the subgroup of male participants. Regarding the local model fit, all estimated parameters were highly significant ($p < .01$).

All fit indices for model fit and model comparisons in the multi-group analyses are displayed in Table 3. Because of the sensitivity of the Chi-Square difference test in large sample sizes (Bentler and Bonett 1980; Meade 2008), we used the changes in the CFI as well as the RMSEA to compare the nested models. For both subgroups (gender and nationality) all successively restricted nested models showed a good model fit. The change in the goodness-of-fit indices remained below the applied cut off values ($\Delta CFI \geq -.01$; Cheung and Rensvold 2002; $\Delta RMSEA \geq .015$; Chen 2007) for all model comparisons. This indicates that the strict or full uniqueness model (M4) was achieved. Thus, the results suggest that all item loadings, intercepts, and residual variances were equivalent across gender and nationality.

4.2 Descriptive Statistics, Reliability and Convergent Validity

Descriptive statistics and reliability coefficients for the German PERMA-Profiler, its subscales and the two validation questionnaires (PWBS and DASS) are presented in Table 4 showing the sample size, number of items per subscale, mean, median, standard deviation, minimum, maximum, skewness, kurtosis and the two reliability indices. The skewness values revealed that all items are negatively skewed. The index for acceptable limits of skewness and kurtosis of ± 2 (Field 2009) showed that all items are slightly to moderately negatively skewed but there are no extreme outliers in our sample ($\text{skewness}_{\max} = -1.27$, $\text{kurtosis}_{\max} = 1.69$; for the detailed descriptive statistics of the items see Appendix Table 8).

Table 3 Model fit and model comparisons testing for measurement invariance of the Five-Factor Model regarding gender and nationality

Model	Model fit					Model comparisons				
	χ^2_{MLR}	df	Scaling	RMSEA	CFI	$\Delta\chi^2_{\text{MLR}}$	Δdf	p	$\Delta RMSEA$	ΔCFI
Gender										
M1: Configural	338.64 **	160	1.272	.051	.967	—	—	—	—	—
M2: Weak	341.58 **	170	1.282	.049	.968	4.96	10	.895	-.002	.001
M3: Strong	376.21 **	180	1.261	.051	.964	40.16	10	.001	.002	-.004
M4: Strict	393.71 **	195	1.285	.049	.963	20.05	15	.087	-.002	-.001
Nationality										
M1: Configural	301.76 **	160	1.337	.046	.972	—	—	—	—	—
M2: Weak	311.66 **	170	1.341	.045	.972	10.30	10	.415	-.001	.000
M3: Strong	334.13 **	180	1.322	.045	.969	23.82	10	.009	.000	-.003
M4: Strict	341.46 **	195	1.343	.042	.971	10.59	15	.781	-.003	.002

** $p < .01$; χ^2_{MLR} - Scaled using maximum likelihood estimation with robust standard errors (MLR); df - Degrees of freedom; Scaling - Scaling correction factor for the Chi-Square from the MLR estimator; $RMSEA$, Root Mean Square Error of Approximation; CFI , Comparative Fit Index; $\Delta\chi^2$, Change in Chi-Square values computed with the Satorra and Bentler (2001) formula; Δdf , Change in degrees of freedom; $\Delta RMSEA$, Change in RMSEA; ΔCFI , Change in CFI

Table 4 Descriptive statistics and reliability coefficients of the PERMA-Profiler, the PWBS, and the DASS

	Items	<i>M</i>	<i>Mdn</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>	<i>Skew</i>	<i>Kurt</i>	λ_6	α
Overall well-being	15	6.98	7.27	1.46	1.40	9.73	-0.88	0.46	.93	.92
Positive Emotions	3	6.57	7.00	1.90	0.00	10.00	-0.83	0.00	.86	.90
Engagement	3	7.12	7.33	1.60	1.00	10.00	-0.93	0.93	.60	.68
Relationships	3	7.34	7.67	2.00	0.00	10.00	-.088	0.29	.71	.78
Meaning	3	6.76	7.00	2.07	0.00	10.00	-0.87	0.30	.83	.88
Accomplishment	3	7.10	7.33	1.56	1.00	10.00	-0.90	0.72	.73	.79
Negative Emotions	3	4.08	4.00	1.86	0.33	9.67	0.31	-0.63	.60	.68
Health	3	6.77	7.33	2.25	0.00	10.00	-0.78	-0.13	.88	.91
PWBS	18	77.30	78.00	9.49	39.00	99.00	-0.62	0.25	.85	.81
DASS	42	26.46	21.00	20.60	0.00	111.00	1.33	1.66	.96	.96

Items - Number of items; *M*, Mean; *Mdn*, Median; *SD*, Standard deviation; *Min*, Minimum; *Max*, Maximum; *Skew*, Skewness; *Kurt*, Kurtosis; λ_6 , Guttman's lambda₆; α , Cronbach's alpha

Considering the internal consistency of the PERMA-Profiler, all five PERMA subscales and the additional domains Health and Negative Emotions showed high reliability coefficients. When interpreting, the low item number ($n = 3$) for each subscale should be considered. The PERMA subscale Engagement ($\lambda_6 = .60$, $\alpha = .68$) showed the lowest, whereas Positive Emotions ($\lambda_6 = .86$, $\alpha = .90$) showed the highest reliability. The overall well-being scale (composed of the main 15 PERMA items) showed excellent reliability values ($\lambda_6 = .94$, $\alpha = .93$). The reliability for the validation questionnaires was sufficiently high for both the PWBS ($\lambda_6 = .85$, $\alpha = .81$) and the DASS ($\lambda_6 = .96$, $\alpha = .96$).

In regard to the convergent validity of the questionnaire, all subscales of the PERMA-Profiler showed consistently positive correlations with the PWBS ($r = .45$ to $r = .77$). Furthermore the overall score of the DASS correlated negatively with all PERMA dimensions ($r = -.58$ to $r = -.70$). Negative Emotions showed negative correlations with all five PERMA dimensions ($r = -.45$ to $r = -.77$). For detailed correlations between the PERMA-Profiler and the different subscales of the PWBS and the DASS see Appendix Table 9.

5 Discussion

The main purpose of the present study was to evaluate the validity of the PERMA theory of well-being by developing and testing a translated and culturally adapted version of the PERMA-Profiler in a large German speaking sample. Because of inconsistencies considering the empirical data relating to the PERMA theory, the study aimed to test which of four theoretically competing models describes best the collected data in German speaking countries and whether PERMA is a uni- or multidimensional construct. The German version of the PERMA-Profiler, which was developed according to international guidelines for cultural adaption (Beaton et al. 2000), revealed high reliability and a high convergent validity.

The CFA results confirmed that the Five-Factor Model showed a better model fit than the Single-Factor and the Higher-Order Model and a nearly equally good fit compared to the Bi-Factor Model. We rejected the Single-Factor Model, because it did not meet the cut-off criteria and showed a considerably inferior model-fit than the multidimensional models. The rejection

of the Single-Factor Model supports the hypothesis that PERMA represents a multidimensional structure. This is in line with the hypothesis of Seligman (2018) that PERMA constitutes the elements of well-being but is not redundant to SWB. Therefore we chose the Five-Factor Model to be the model which shows the best trade-off between statistical model-fit and theoretical interpretability. To our knowledge, this is the first study to suggest the multidimensional structure of the PERMA-Profil by comparing four different theoretically possible factor solutions. We tested for measurement invariance considering two subgroups for gender (male, female) and for nationality (Austrian, German), respectively. The CFA results confirmed the Five-Factor Model within both subgroups. Multi-group CFA results revealed measurement invariance across gender and nationality. A detailed discussion of these major findings is provided below.

The German version of the PERMA-Profil showed excellent reliability coefficients for the overall well-being scale. Considering the subscales, Positive Emotions and Meaning showed good, Relationships and Accomplishment showed acceptable reliability. Only the subscale Engagement revealed a Cronbach's alpha of $\alpha = .68$, which can be contributed to the contentual heterogeneity of this domain combined with the brief structure of the PERMA-Profil (i.e., three items per domain). Internal consistency indices, such as Cronbach's alpha, rise with the number of items used (Tavakol and Dennick 2011). From this perspective the low values considering Engagement can be seen as trade-off considering the shortness of the questionnaire.

Considering the validity indices of the German version of the PERMA-Profil, we could expand the findings of Butler and Kern (2016) through the inclusion of two measures, which were not included in the validation study of the English original (PWBS; Ryff and Keyes 1995; DASS; Nilges and Essau 2015). All PERMA subscales showed a positive correlation with the PWBS scores. Compared to the initial findings of Butler and Kern (2016), the correlations of the German version subscales with the PWBS were equally high as were the correlations of the English original subscales with the Satisfaction with Life Scale (Diener et al. 1985). Considering the convergent validity, all subscales of the PERMA-Profil showed consistently negative correlations with the overall score of the DASS, which includes negative emotional states such as depression, anxiety and stress. Very similar correlation coefficients with related constructs could be found in the publication of Butler and Kern (2016). These results support the hypothesis that the validity PERMA-Profil is stable for English and German speaking populations, respectively.

In comparison with theoretically competing model constellations, the inter-correlated Five-Factor Model turned out to have a better global model fit than a Single-Factor Model and a Higher-Order Model, disclosed through the values of the goodness-of-fit indices. The rejection of the Single-Factor Model advocates strongly for the multidimensional nature of the construct. The high latent inter-correlation between the PERMA domains ($r = .53-.82$) are corrected for attenuation in the CFA model and in line with the results of the studies presented by Butler and Kern. The high correlations suggest the interdependent nature of the domains, as well as the far better fit of the multidimensional model speaks for the existence of the distinct domains. Compared to the Bi-Factor Model, the Five-Factor Model showed no distinct difference considering the model-fit indices. We rejected the Bi-Factor Model because some of its strict assumptions do not match well with the original assumptions of the PERMA-theory. The Bi-Factor model assumes one underlying general well-being factor, which is directly connected to the item responses. Thus, in this model the residual PERMA factors would only explain additional variance and are not defined as the postulated "building blocks"

of SWB. The Bi-Factor model suppresses the strong latent inter-correlations between the residual PERMA facets. This would lead to serious difficulties in interpreting the results of the PERMA-questionnaire, which emanate from the assumption that the single PERMA-domains are interdependent and promote each other. The Five-Factor Model supports the hypothesis of PERMA as multidimensional construct, which is not redundant to SWB. This is in line with the postulates of the original PERMA theory by Seligman (2011, 2018). Our results indicate that the Five-Factor-Model, which is also psychometrically tested in multiple English speaking populations (Butler and Kern 2016) is also well-suited to operationalize SWB in German speaking populations.

The PERMA-theory can be seen as expansion of the original theory of SWB that provides an insight in the components which lead to the experience of SWB. Initial data was provided by the original study of Butler and Kern (2016), which found correlations between SWB and the PERMA components between $r = .44$ and $r = .76$. In our study we found inter-correlations between $r = .45$ and $r = .70$ between the PERMA domains and the PWBS, which can be seen as comparably high as in the study of Butler and Kern. These findings support the assumption of Seligman (2018), that PERMA is not an entirely new form of well-being, but provides information about the individual composition of the individual overall well-being. Psychometrically seen the PERMA theory as well as the corresponding questionnaire is not as parsimonious as SWB, but it provides valuable information for the development of specific interventions for the promotion of one PERMA component in the context of psychotherapy or coaching (Seligman et al. 2006; For an overview on the effectiveness of Positive Interventions see Bolier et al. 2013).

To examine the hypotheses that the Five-Factor Model is invariant across gender and nationality, we calculated multi-group CFAs. In the first step, single group analyses revealed acceptable fit indices for subgroups according to gender (female vs. male) and nationality (Austrian vs. German). The only exception was the RMSEA in the male subgroup ($n = 140$), which was slightly higher than in the female subgroup. For small sample sizes more liberal cut-off criteria can be accepted (Beauducel and Wittmann 2005; Heene et al. 2011; Hu and Bentler 1998, 1999), which supports the evidence that the Five-Factor Model is acceptable for male and female adults, respectively.

The multi-group CFAs revealed strict or full uniqueness measurement invariance of the Five-Factor Model, meaning equal factor loadings, intercepts and residuals for men and women as well as Austrian and German citizen, respectively. We used widely accepted cut-off values for detecting invariance: $\Delta CFI \geq -.01$ (Cheung and Rensvold 2002) and $\Delta RMSEA \geq .015$ (Chen 2007). Although, Meade et al. (2008) postulated more conservative cut-off values of $\Delta CFI \geq -.002$, the authors state that lower standards should be used for translations and correlation purposes. Little (2013) concludes that the cut-off values of Meade et al. (2008) are too strict for real world problems and that the $\Delta CFI \geq -.01$ is still acceptable. Eventually, there is great consensus among the mentioned authors that the Chi-Square difference test is too sensitive for large sample sizes and violations of the normal distribution assumption. To our knowledge, these are the first findings, which support measurement invariance for the PERMA-Profiler considering gender and German speaking nationalities. These results indicate that the scores among men and woman as well as Austrian and German citizens can be compared and interpreted meaningfully.

Despite the promising results, with regard to the successful adaption of the PERMA-Profiler for German speaking countries, there is a need to discuss certain limitations of the study. At first, we need to mention limitations concerning the convenience sample. Even

though the study reached a very large sample size, online sample acquisition led to an imbalance regarding gender, age and nationality. Within the sample, females and young people were overrepresented, whereas the Swiss nationality was hardly reached. Because of the small number of Swiss participants, our results can only confirm the measurement invariance of the Bi-Factor Model between the Austrian and German nationalities. In future studies the measurement invariance of the PERMA theory should work with more representative samples, which include a balanced sample composition regarding gender and nationality.

Another major point for discussion is that all items of the PERMA-Profil were slightly to moderately deviant from normality, which can lead to biased statistical results (Satorra and Bentler 1994). Even so, the observed negatively skewed distribution corresponds highly with the item parameters reported by Butler and Kern (2016). A similar effect can be observed for the results of the Better Life Index of the OECD (2016), which shows a mean of 7 (maximum = 10) for Austria and Germany. A possible explanation could be the phenomenon of social desirability, which can occur in the context of well-being assessments. However, research on this topic showed that the control of social desirability is at the expense of the validity of the questionnaire (Diener et al. 1991). Diener (1994) suggests viewing social desirability as inevitable part of well-being assessments.

This is the first study which supports the validity of the Five-Factor Model of the PERMA theory in German speaking countries. Since, we could not establish a clear psychometrical difference between the Five-Factor-Model and the Bi-Factor-Model, we suggest that future research should focus on the psychometrical comparison between the Five-Factor Model and other model solutions. Until now, this is the second study to examine the validity of the PERMA-Profil after the study of Butler and Kern (2016), who investigated English speaking countries. Further research should be done investigating the validity of the Five-Factor Model of PERMA in different cultures and languages. Future research could determine whether the PERMA factors hold some incremental validity in predicting an outcome such as mental health.

Due to the fact that the PERMA-Profil represents a quite new well-being questionnaire, there is no research we are aware of, on the inter-individual differences of the PERMA factors considering different sociodemographic variables. However, there is research on the gender differences in different other well-being questionnaires (e.g. Al-Attiah and Nasser 2016; Chui and Wong 2016; Diener 1984; Li et al. 2015; Lindfors et al. 2006). Furthermore, research shows differences in well-being considering other variables such as age, culture and socioeconomic status (Meisenberg and Woodley 2015; Pinquart and Sörensen 2000; Tesch-Römer et al. 2008). Our results are the first to demonstrate the measurement invariance for gender and nationality for the PERMA-Profil, which is the psychometric requirement to interpret latent mean differences meaningfully. Further research on the validity of the PERMA theory should also focus on the potential inter-individual differences between the five PERMA factors considering various sociodemographic variables such as different nationalities, socioeconomic status and gender.

Taken together, the development and evaluation of a German version of the PERMA-Profil confirmed the multidimensional structure of the PERMA theory in a German speaking population. The Five-Factor model showed the best trade-off between psychometric model-fit and theoretical interpretability. In addition, this study provides first evidence for the measurement invariance of the PERMA-Profil considering gender and nationality. Further research

should promote the test for measurement invariance for the five PERMA domains considering relevant sociodemographic variables and the test of the Five-Factor Model in other cultures and language regions.

Acknowledgements The contributions of the first two authors (MW & JJ) to this manuscript were equal. Open access funding provided by University of Graz.

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no conflict of interest.

Statement of Human Rights All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed Consent Informed consent was obtained from all individual participants included in the study. The informed consent was part of the first page of the online-survey. Every participant could choose, whether he/she wants to participate in the study or not. No identifying information about participants was collected.

Appendix 1

Table 5 Item labels, German item wordings and scale assignments of the German version of the PERMA-Profiler

Label	Items (in order of appearance)	Scale ^b
A1	Wie oft haben Sie das Gefühl sich der Erreichung Ihrer Ziele zu nähern?	a
E1	Wie oft sind Sie ganz in dem versunken was Sie gerade tun?	a
P1	Wie oft fühlen Sie sich im Allgemeinen voller Freude?	a
N1	Wie oft fühlen Sie sich ängstlich?	a
A2	Wie oft erreichen Sie wichtige selbstgesetzte Ziele?	a
H1	Wie würden Sie Ihre Gesundheit im Allgemeinen einschätzen?	c
M1	In welchem Ausmaß führen Sie ein sinnvolles und bedeutsames Leben?	b
R1	In welchem Ausmaß erhalten Sie Hilfe und Unterstützung von anderen, wenn Sie diese brauchen?	b
M2	In welchem Ausmaß glauben Sie, dass das, was Sie in ihrem Leben tun wertvoll und lohnenswert ist?	b
E2	In welchem Maß können Sie sich für Dinge interessieren und begeistern?	b
Lo	Wie einsam fühlen Sie sich in Ihrem Alltagsleben?	b
H2	Wie zufrieden sind Sie derzeit mit Ihrer Gesundheit?	b
P2	Wie oft fühlen Sie sich im Allgemeinen positiv gestimmt?	a
N2	Wie oft fühlen Sie sich wütend?	a
A3	Wie oft sind Sie in der Lage Ihren Verantwortungen gerecht zu werden?	a
N3	Wie oft fühlen Sie sich traurig?	a
E3	Wie oft vergessen Sie die Zeit während Sie etwas tun, das Sie genießen?	a
H3	Wie schätzen Sie Ihre Gesundheit im Vergleich zu anderen ein, die dasselbe Geschlecht und Alter wie Sie haben?	c
R2	In welchem Ausmaß haben Sie sich bisher geliebt gefühlt?	b
M3	In welchem Ausmaß glauben Sie, dass Sie im Leben eine Orientierung haben?	b
R3	Wie zufrieden sind Sie mit ihren persönlichen Beziehungen?	b
P3	In welchem Ausmaß fühlen Sie sich zufrieden?	b
Ha	Wie glücklich fühlen Sie sich, wenn sie alle Aspekte Ihres Lebens betrachten?	b

^a First letter refers to the PERMA domain, the number denotes the chronology within the domain; ^b Rating scale anchors: a - never - always, b - not at all - completely, c - very badly - excellent; P, Positive Emotions; E, Engagement; R, Positive Relationships; M, Meaning; A, Accomplishment; N, Negative Emotions; H, Health; Lo, Loneliness; Ha, Overall happiness

Appendix 2

Table 6 Standardized factor loadings for the Bi-Factor Model using the MLR estimator

Scale	Item	Overall well-being	P	E	R	M	A
Positive Emotions	1	.79	.35				
	2	.79	.50				
	3	.85	.20				
Engagement	1	.48		.70			
	2	.52		.28			
	3	.34		.47			
Positive Relationships	1	.55			.41		
	2	.58			.42		
	3	.65			.50		
Meaning	1	.79				.38	
	2	.72				.55	
	3	.70				.30	
Accomplishment	1	.74					.37
	2	.63					.59
	3	.53					.26

MLR, Maximum likelihood estimation with robust standard errors; *P*, Positive Emotions; *E*, Engagement; *R*, Positive Relationships; *M*, Meaning; *A*, Accomplishment

Appendix 3

Table 7 Standardized factor loadings for the Single-Factor Model and the Five-Factor Model using the MLR Estimator

Scale	Item	Single-factor model: Overall well-being	Correlated five-factor model				
			P	E	R	M	A
Positive emotions	1	.81	.86				
	2	.81	.87				
	3	.85	.88				
Engagement	1	.49		.74			
	2	.53		.64			
	3	.35		.56			
Positive relationships	1	.56			.69		
	2	.60			.72		
	3	.66			.81		
Meaning	1	.80				.89	
	2	.75				.86	
	3	.72				.77	
Accomplishment	1	.75					.87
	2	.65					.77
	3	.54					.60
Latent factor correlations							
Positive emotions				.64	.82	.79	.75
Engagement					.53	.61	.60
Relationships						.65	.60
Accomplishment							.81

MLR, Maximum likelihood estimation with robust standard errors; *P*, Positive Emotions; *E*, Engagement; *R*, Positive Relationships; *M*, Meaning; *A*, Accomplishment

Appendix 4

Table 8 Descriptive statistics of the German version of the PERMA-Profiler

	Item	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>	<i>Skew</i>	<i>Kurt</i>
Positive emotions	1	6.26	2.09	0	10	−0.76	−0.19
	2	6.78	1.93	1	10	−0.87	0.21
	3	6.66	2.22	0	10	−0.87	0.21
Engagement	1	6.29	2.05	0	10	−0.76	0.23
	2	8.03	1.82	0	10	−1.20	1.69
	3	7.03	2.25	0	10	−0.93	0.37
Positive relationships	1	7.76	2.23	0	10	−1.27	1.31
	2	7.30	2.49	0	10	−0.86	−0.16
	3	6.95	2.31	0	10	−0.86	0.18
Meaning	1	6.61	2.30	0	10	−0.77	0.03
	2	6.85	2.28	0	10	−0.88	0.37
	3	6.84	2.36	0	10	−0.90	0.26
Accomplishment	1	6.58	1.95	0	10	−0.77	0.42
	2	7.23	1.90	0	10	−1.08	0.94
	3	7.48	1.71	0	10	−0.98	1.04
Negative emotions	1	4.07	2.60	0	10	0.32	−1.08
	2	3.88	2.28	0	10	0.47	−0.77
	3	4.30	2.24	0	10	0.33	−0.94
Health	1	6.98	2.21	0	10	−0.91	0.31
	2	6.76	2.65	0	10	−0.79	0.08
	3	6.55	2.45	0	10	−0.66	−0.33
Loneliness		3.55	2.76	0	10	0.47	−0.96
Overall happiness		6.89	2.17	0	10	−0.96	0.51

M, Mean; *Mdn*, Median; *SD*, Standard deviation; *Min*, Minimum Value; *Max*, Maximum value; *Skew*, Skewness; *Kurt*, Kurtosis

Appendix 5

Table 9 Inter-correlations of the PERMA-Profiler and correlations with the validation scales

	Items	P	E	R	M	A	Overall	N	H
Positive emotions	3								
Engagement	3	.53**							
Relationships	3	.69**	.39**						
Meaning	3	.70**	.46**	.54**					
Accomplishment	3	.64**	.43**	.48**	.70**				
Overall well-being	15	.89**	.68**	.79**	.86**	.80**			
Negative emotions	3	−.59**	−.23**	−.42**	−.45**	−.42**	−.53**		
Health	3	.50**	.30**	.38**	.38**	.38**	.47**	.39**	
Loneliness	1	−.58**	−.25**	−.58**	−.47**	−.43**	−.59**	.52**	−.25**
Overall happiness	1	.81**	.44**	.66**	.69**	.61**	.81**	−.47**	.47**
PWBS	18	.68**	.45**	.61**	.70**	.65**	.77**	−.54**	.40**
Autonomy	3	.25**	.30**	.17**	.30**	.27**	.32**	−.28**	.15**
Environmental Mastery	3	.67**	.36**	.49**	.59**	.61**	.68**	−.62**	.46**
Personal growth	3	.36**	.36**	.28**	.39**	.37**	.44**	−.26**	.26**
Positive Relationships	3	.53**	.28**	.64**	.44**	.37**	.57**	−.38**	.29**
Self-acceptance	3	.73**	.42**	.57**	.73**	.66**	.78**	−.53**	.37**
Purpose in life	3	−.03**	−.03**	.08**	.14**	.15**	.08**	.08**	.02**

Table 9 (continued)

	Items	P	E	R	M	A	Overall	N	H
DASS	42	-.69**	-.36**	-.51**	-.56**	-.51**	-.66**	.70**	-.45**
Depression	18	-.75**	-.42**	-.59**	-.66**	-.58**	-.76**	.61**	-.40**
Anxiety	18	-.49**	-.26**	-.35**	-.40**	-.38**	-.47**	.58**	-.41**
Stress	18	-.54**	-.24**	-.36**	-.38**	-.34**	-.47**	.66**	-.37**

$N = 854$ for all correlations; ** $p < .01$, * $p < .05$; *P*, Positive emotions; *E*, Engagement; *R*, Positive relationships; *M*, Meaning; *A*, Accomplishment; *N*, Negative emotions; *H*, Health; *Overall*, general well-being score (i.e. average of the PERMA 15 items); *PWBS*, Psychological Well-Being Scale; *DASS*, Depression Anxiety Stress Scale

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Abbott, R. A., Ploubidis, G. B., Huppert, F. A., Kuh, D., Wadsworth, M. E. J., & Croudace, T. J. (2006). Psychometric evaluation and predictive validity of Ryff's psychological wellbeing items in a UK cohort sample of women. *Health and Quality of Life Outcomes*, 4, 76. <https://doi.org/10.1186/1477-7525-4-76>.
- Al-Attayah, A., & Nasser, R. (2016). Gender and age differences in life satisfaction within a sex-segregated society: Sampling youth in Qatar. *International Journal of Adolescence and Youth*, 21(1), 84–95. <https://doi.org/10.1080/02673843.2013.808158>.
- Ammon, U. (1995). *Die deutsche Sprache in Deutschland, Österreich und der Schweiz. Das Problem der nationalen Varietäten*. Berlin/New York: de Gruyter. <https://doi.org/10.1515/9783110872170>.
- Bartkowiak, A. (2008). *Diplomarbeit Eine umfassende Betrachtung der Einflussfaktoren Zusammenhänge - unter Einbeziehung religiöser und spiritueller Variablen*. Universität Wien.
- Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine*, 25(24), 3186–3191. <https://doi.org/10.1097/00007632-200012150-00014>.
- Beauducel, A., & Wittmann, W. W. (2005). Simulation study on fit indexes in CFA based on data with slightly distorted simple structure. *Structural Equation Modeling: A Multidisciplinary Journal*, 12(1), 41–75. https://doi.org/10.1207/s15328007sem1201_3.
- Beaujean, A. (2015). John Carroll's views on intelligence: Bi-factor vs. higher-order models. *Journal of Intelligence*, 3, 121–136. <https://doi.org/10.3390/jintelligence3040121>.
- Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, 88(3), 588–606. <https://doi.org/10.1037/0033-2909.88.3.588>.
- Bolier, L., Haverman, M., Westerhof, G. J., Riper, H., Smit, F., & Bohlmeijer, E. (2013). Positive psychology interventions: A meta-analysis of randomized controlled studies. *BMC Public Health*, 13(1), 13–119.
- Bradburn, N. M. (1969). *The structure of psychological well-being*. Chicago: Aldine.
- Butler, J., & Kern, M. L. (2016). The PERMA-profiler: A brief multidimensional measure of flourishing. *International Journal of Wellbeing*, 6, 1–48. <https://doi.org/10.5502/ijw.v6i3.1>.
- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 14(3), 464–504. <https://doi.org/10.1080/10705510701301834>.
- 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. <https://doi.org/10.1007/s10902-012-9367-x>.
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, 9(2), 233–255. <https://doi.org/10.1207/S15328007SEM0902>.
- Chui, W. H., & Wong, M. Y. H. (2016). Gender differences in happiness and life satisfaction among adolescents in Hong Kong: Relationships and self-concept. *Social Indicators Research*, 125(3), 1035–1051. <https://doi.org/10.1007/s11205-015-0867-z>.

- Coffey, J. K., Wray-Lake, L., Mashek, D., & Branand, B. (2014). A multi-study examination of well-being theory in college and community samples. *Journal of Happiness Studies*, 17(1), 187–211. <https://doi.org/10.1007/s10902-014-9590-8>.
- De Bruin, G. P., & Du Plessis, G. A. (2015). Bifactor analysis of the mental health continuum-short form (MHC-SF). *Psychological Reports*, 116, 438–446. <https://doi.org/10.2466/03.02.PR0.116k20w6>.
- Delle Fave, A., Massimini, F., & Bassi, M. (2011). *Psychological selection and optimal experience across cultures. Cross-cultural advancements in positive psychology* (Vol. 2). Dordrecht: Springer Netherlands. <https://doi.org/10.1007/978-90-481-9876-4>.
- Diener, E. (1984). Subjective well-being. *Psychological Bulletin*, 95(3), 542–575. <https://doi.org/10.1037/0033-2909.95.3.542>.
- Diener, E. (1994). Assessing subjective well-being: Progress and opportunities. *Social Indicators Research*, 31(2), 103–157. <https://doi.org/10.1007/BF01207052>.
- Diener, E., & Ryan, K. (2009). Subjective well-being: A general overview. *South Africa Journal of Psychology*, 39(4), 391–406. <https://doi.org/10.1177/008124630903900402>.
- Diener, E., & Suh, E. (1997). Measuring quality of life: Economic, social, and subjective indicators. *Social Indicators Research*, 40(1/2), 189–216. <https://doi.org/10.1023/A:1006859511756>.
- Diener, E., Emmons, R., Larsen, R., & Griffin, S. (1985). The satisfaction with life scale. *Journal of Personality Assessment*, 49(1), 71–75. https://doi.org/10.1207/s15327752jpa4901_13.
- Diener, E., Sandvik, E., Pavot, W., & Gallagher, D. (1991). Response artifacts in the measurement of subjective well-being. *Social Indicators Research*, 24(1), 35–56. <https://doi.org/10.1007/BF00292649>.
- Diener, E., Wirtz, D., Biswas-Diener, R., Tov, W., Kim-Prieto, C., Choi, D., & Oishi, S. (2009). New measures of well-being. In E. Diener (Ed.), *Assessing well-being*, 39 (pp. 247–266). Dordrecht: Springer. <https://doi.org/10.1007/978-90-481-2354-4>.
- Dodge, R., Daly, A. P., Huyton, J., & Sanders, L. D. (2012). The challenge of defining wellbeing. *International Journal of Wellbeing*, 2, 222–235. <https://doi.org/10.5502/ijw.v2i3.4>.
- Dodou, D., & de Winter, J. C. F. (2014). Social desirability is the same in offline, online, and paper surveys: A meta-analysis. *Computers in Human Behavior*, 36, 487–495. <https://doi.org/10.1016/j.chb.2014.04.005>.
- Dunbar, M., Ford, G., Hunt, K., & Der, G. (2000). Question wording effects in the assessment of global self-esteem. *European Journal of Psychological Assessment*, 16(1), 13–19. <https://doi.org/10.1027//1015-5759.16.1.13>.
- Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review*, 87(3), 215–251. <https://doi.org/10.1037//0033-295X.87.3.215>.
- Field, A. (2009). *Discovering statistics using SPSS* (3rd ed.). London: Sage Publications Ltd.
- Finney, S. J., & DiStefano, C. (2013). Non-normal and categorical data in structural equation modeling. In G. R. Hancock & O. Mueller (Eds.), *Quantitative methods in education and the behavioral sciences: Issues, research, and teaching. Structural equation modeling: A second course* (2nd ed., pp. 439–492). Charlotte, NC, US.
- Goodman, F. R., Disabato, D. J., Kashdan, T. B., & Kauffman, S. B. (2017). Measuring well-being: A comparison of subjective well-being and PERMA. *The Journal of Positive Psychology*, 13(4), 1–12. <https://doi.org/10.1080/17439760.2017.1388434>.
- Heene, M., Hilbert, S., Draxler, C., Ziegler, M., & Bühner, M. (2011). Masking misfit in confirmatory factor analysis by increasing unique variances: A cautionary note on the usefulness of cutoff values of fit indices. *Psychological Methods*, 16(3), 319–336. <https://doi.org/10.1037/a0024917>.
- Henderson, L., & Knight, T. (2012). Integrating the hedonic and eudaimonic perspectives to more comprehensively understand wellbeing and pathways to wellbeing. *International Journal of Wellbeing*, 2(3), 196–221. <https://doi.org/10.5502/ijw.v2i3.3>.
- Hides, L., Quinn, C., Stoyanov, S., Cockshaw, W., Mitchell, T., & Kavanagh, D. J. (2016). Is the mental wellbeing of young Australians best represented by a single, multidimensional or bifactor model? *Psychiatry Research*, 241, 1–7. <https://doi.org/10.1016/j.psychres.2016.04.077>.
- Hu, L., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, 3(4), 424–453. <https://doi.org/10.1037//1082-989X.3.4.424>.
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>.
- Huppert, F. A., Marks, N., Clark, A., Siegrist, J., Stutzer, A., Vittersø, J., & Wahrendorf, M. (2009). Measuring well-being across Europe: Description of the ESS well-being module and preliminary findings. *Social Indicators Research*, 91(3), 301–315. <https://doi.org/10.1007/s11205-008-9346-0>.
- 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. <https://doi.org/10.1016/j.paid.2014.11.026>.

- Kashdan, T. B. (2017, October 12). How many ways can we measure well-being? Reconsidering how we think about the subject. *Psychology Today*. Retrieved from <https://www.psychologytoday.com/intl/blog/curious/201710/how-many-ways-can-we-measure-well-being>. Accessed 12 Oct 2017.
- Keyes, C. L., Wissing, M., Potgieter, J. P., Temane, M., Kruger, A., & van Rooy, S. (2008). Evaluation of the mental health continuum-short form (MHC-SF) in Setswana-speaking south Africans. *Clinical Psychology & Psychotherapy*, 15, 181–192. <https://doi.org/10.1002/cpp.572>.
- Kline, R. B. (2015). *Principles and practice of structural equation modelling* (4th ed.). New York: Guilford Press.
- Krueger, J. I. (2012). Seligman's flourish: The second coming. *The American Journal of Psychology*, 125(1), 121–124. <https://doi.org/10.5406/amerjpsyc.125.1.0121>.
- Lamborn, P., Cramer, K. M., & Riberdy, A. (2018). The structural validity and measurement invariance of the mental health continuum – short form (MHC-SF) in a large Canadian sample. *Journal of Well-being Assessment*, 2(1), 1–19. <https://doi.org/10.1007/s41543-018-0007-z>.
- Layard, R. (2010). Measuring subjective well-being. *Science*, 327(5965), 534–535. <https://doi.org/10.1126/science.1186315>.
- Li, R.-H., Kao, C.-M., & Wu, Y.-Y. (2015). Gender differences in psychological well-being: Tests of factorial invariance. *Quality of Life Research*, 24(11), 2577–2581. <https://doi.org/10.1007/s11136-015-0999-2>.
- LimeSurvey Project Team. (2015). LimeSurvey: An open source survey tool. Retrieved from <http://www.limesurvey.org>
- Lindfors, P., Berntsson, L., & Lundberg, U. (2006). Total workload as related to psychological well-being and symptoms in full-time employed female and male white-collar workers. *International Journal of Behavioral Medicine*, 13(2), 131–137. https://doi.org/10.1207/s15327558ijbm1302_4.
- Little, T. D. (2013). *Longitudinal structural equation modeling*. New York: Guilford Press.
- Lovibond, P. F., & Lovibond, S. H. (1995). The structure of negative emotional states: Comparison of the depression anxiety and stress scales (DASS) with the Beck depression and anxiety inventories. *Behaviour Research and Therapy*, 33(3), 335–343.
- Meade, A. W. (2008, April). *Power of AFI to detect CFA model misfit*. Paper presented at the 23rd Annual Conference of the Society for Industrial and Organizational Psychology. San Francisco, CA.
- Meade, A. W., Johnson, E. C., & Braddy, P. W. (2008). Power and sensitivity of alternative fit indices in tests of measurement invariance. *The Journal of Applied Psychology*, 93(3), 568–592. <https://doi.org/10.5465/AMBPP.2006.27182124>.
- Meisenberg, G., & Woodley, M. A. (2015). Gender differences in subjective well-being and their relationships with gender equality. *Journal of Happiness Studies*, 16(6), 1539–1555. <https://doi.org/10.1007/s10902-014-9577-5>.
- Nilges, P., & Essau, C. (2015). Die Depressions-Angst-Stress-Skalen: Der DASS – ein Screeningverfahren nicht nur für Schmerzpatienten. *Der Schmerz*, 10(4), 649–657. <https://doi.org/10.1007/s00482-015-0019-z>.
- Pinquart, M., & Sörensen, S. (2000). Influences of socioeconomic status, social network, and competence on subjective well-being in later life: A meta-analysis. *Psychology and Aging*, 15(2), 187–224. <https://doi.org/10.1037/0882-7974.15.2.187>.
- Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological Methodology*, 25, 111–163.
- Reise, S. P. (2012). The rediscovery of bifactor measurement models. *Multivariate Behavioral Research*, 47, 667–696. <https://doi.org/10.1080/00273171.2012.715555>.
- Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48(2), 1–36. doi:<https://doi.org/10.18637/jss.v048.i02>.
- RStudio (2016). RStudio: Integrated development for R (version 1.1.456) [computer software]. Boston. Retrieved from: <http://www.rstudio.com/>
- 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.
- Ryff, C. D., & Keyes, C. L. (1995). The structure of psychological well-being revisited. *Journal of Personality and Social Psychology*, 69(4), 719–727. <https://doi.org/10.1037/0022-3514.69.4.719>.
- Satorra, A., & Bentler, P. M. (1994). Corrections to test statistics and standard errors in covariance structure analysis. In A. Von Eye, C. Clog, & Clifford (Eds.), *Latent variables analysis: Applications to developmental research* (pp. 399–419). Thousand Oaks: SAGE Publications, Inc.
- Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*, 66(4), 507–514. <https://doi.org/10.1007/BF02296192>.
- Schutte, L., & Wissing, M. P. (2017). Clarifying the factor structure of the mental health continuum short form in three languages: A Bifactor exploratory structural equation modeling approach. *Society and Mental Health*, 7(3), 142–158. <https://doi.org/10.1177/2156869317707793>.
- Seligman, M. E. P. (2002). *Authentic happiness: Using the new positive psychology to realize your potential for lasting fulfillment*. New York: Simon and Schuster.

- Seligman, M. E. P. (2011). *Flourish: A visionary new understanding of happiness and well-being*. New York, NY, US: Free Press.
- Seligman, M. E. P. (2018). PERMA and the building blocks of well-being. *The Journal of Positive Psychology*, 13(4), 333–335. <https://doi.org/10.1080/17439760.2018.1437466>.
- Seligman, M. E. P., Rashid, T., & Parks, A. C. (2006). Positive psychotherapy. *American Psychologist*, 61(8), 774–788.
- Shin, D., & Johnson, D. (1978). Avowed happiness as an overall assessment of the quality of life. *Social Indicators Research*, 5(1–4), 475–492.
- Silverman, A. L., Forgeard, M., Beard, C., & Björgvinsson, T. (2018). Psychometric properties of the mental health continuum – short form in a psychiatric sample. *Journal of Well-being Assessment*, 2(1), 57–73. <https://doi.org/10.1007/s41543-018-0011-3>.
- Steiger, J. H., Shapiro, A., & Browne, M. W. (1985). On the multivariate asymptotic distribution of sequential Chi-square statistics. *Psychometrika*, 50(3), 253–263. <https://doi.org/10.1007/BF02294104>.
- Stoll, L. (2014). A short history of wellbeing research. *Wellbeing*, 5(1), 1–19. <https://doi.org/10.1002/9781118539415.WBWELL098>.
- Su, R., Tay, L., & Diener, E. (2014). The development and validation of the comprehensive inventory of thriving (CIT) and the brief inventory of thriving (BIT). *Applied Psychology: Health and Well-Being*, 6(3), 251–279. <https://doi.org/10.1111/aphw.12027>.
- Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, 2, 53–55. <https://doi.org/10.5116/ijme.4dfb.8dfd>.
- Tesch-Römer, C., Motel-Klingebiel, A., & Tomasik, M. J. (2008). Gender differences in subjective well-being: Comparing societies with respect to gender equality. *Social Indicators Research*, 85(2), 329–349. <https://doi.org/10.1007/s11205-007-9133-3>.
- 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(1), 4–70.
- Van Zyl, L. E. (2013). Seligman's flourishing: An appraisal of what lies beyond happiness. *SA Journal of Industrial Psychology*, 39(2), 3 pages. <https://doi.org/10.4102/sajip.v39i2.1168>.
- Wilson, W. R. (1967). Correlates of avowed happiness. *Psychological Bulletin*, 67(4), 294–306.
- Wolgast, A., Stiensmeier-Pelster, J., & von Aufschnaiter, C. (2014). Papierbasierte oder internetbasierte Skalen zur Erfassung von Motivation (SELLMO) und Selbstkonzept (SESSKO)? *Diagnostica*, 60(1), 46–58. <https://doi.org/10.1026/0012-1924/a000100>.
- Yuan, K. H., & Bentler, P. M. (2000). Three likelihood-based methods for mean and covariance structure analysis with nonnormal missing data. *Sociological Methodology*, 30(1), 165–200. <https://doi.org/10.1111/0081-1750.00078>.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.