J Happiness Stud
DOI 10.1007/s10902-014-9590-8

RESEARCH PAPER

A Multi-Study Examination of Well-Being Theory in College and Community Samples

John K. Coffey · Laura Wray-Lake · Debra Mashek · Brittany Branand

Abstract Well-being theory (WBT) proposes five indicators of well-being [i.e., positive emotion, engagement, relationships, meaning, achievement (PERMA)] that are, independently, empirically supported predictors of flourishing (i.e., an optimal level of well-being; Seligman in Flourish: a visionary new understanding of happiness and well-being. Free Press, NY, 2011). However, there is limited empirical support for the multidimensional model suggested by WBT. Two studies sought to test and validate the higher-order factor structure of the five components of PERMA and PERMA's ability to predict concurrent and prospective flourishing outcomes (e.g., physical health, college success). In Study 1, a longitudinal examination of college students, participants completed measures of well-being (including four of the five PERMA indicators), physical health, and college success at the end of their sophomore, junior, and senior years. In Study 2, a larger, cross-sectional study was conducted online to further validate the PERMA model with a broader sample and all five PERMA indicators. Participants completed measures similar to those administered at Study 1 and other measures used to validate Study 1 measures. Results from Study 2 further validated the PERMA model by comparing Study 1 measures to established measures and by adding meaning to the model. Study 1 and Study 2 PERMA models predicted markers of well-being (e.g., vitality, life satisfaction) and flourishing (e.g., physical health). The two studies reported here provide cross-sectional and longitudinal support that WBT is useful for predicting flourishing.

J. K. Coffey (✉) · B. Branand
Division of Behavioral and Organizational Sciences, Claremont Graduate University, 123 E. Eighth Street, Claremont, CA 91711, USA
e-mail: john.coffey@cgu.edu

L. Wray-Lake
Department of Clinical and Social Sciences in Psychology, University of Rochester, Rochester, NY, USA

D. Mashek
Department of Humanities, Social Sciences, and Arts, Harvey Mudd College, Claremont, CA, USA

Published online: 29 October 2014

Keywords Well-being theory · PERMA · Engagement · Relationships achievement · Health

1 Introduction

Around the world, people pursue the "good life" (Diener 2000; Seligman and Csikszentmihalyi 2000), and they do so in varied ways. One person might seek wealth and fame; another might seek authentic relationships and positive change in the community. This "good life," as represented by high levels of well-being or flourishing (an optimal level of well-being that extends beyond the mere absence of pathologies; Huppert 2009; Keyes 2007; Seligman 2011), is within grasp; yet, at any given point in time, less than 20 % of us are living the good life (Keyes 2007). This aggregate level of sub-optimal functioning is unfortunate because, on the whole, well-being affords a host of beneficial outcomes for individuals (e.g., better health, stronger relationships) and society (e.g., greater work productivity, more prosocial behavior; Diener et al. 2010; Huppert 2009; Keyes 2007; Lyubomirsky et al. 2005). Understanding the factors that move a person closer to flourishing is an obvious next step toward helping more of us live the good life. Using a college student sample and broader adult sample, this manuscript examines a promising new theory, Seligman's multidimensional well-being theory (WBT) (Seligman 2011), by longitudinally testing and validating the higher-order factor structure of the proposed five-dimensional model of well-being. Then, this manuscript tests the ability of the higher-order factor to predict flourishing (e.g., physical health, success) in these two populations.

The question of what exactly promotes well-being is an open one that has compelled a good deal of theoretical and empirical work. Much of the early work took a unidimensional approach in which scholars sought to identify individual factors (e.g., positive emotions, engagement, meaning in life) that were key indicators of well-being (e.g., Csikszentmihalyi 2008; Fredrickson 2001; Steger et al. 2009).

With the benefit of the strong foundation created by unidimensional approaches to well-being (described further below), and recognizing that people are complex creatures who pursue the good life in a multitude of ways, theoretical and empirical integration became a disciplinary priority. Multidimensional theories of well-being sought to understand how different factors work in concert to give rise to the good life we all seek (e.g., Diener et al. 2010; Huppert and So 2013; Keyes 2007; Ryan and Deci 2001; Ryff 1989; Seligman 2011).

One promising multidimensional theory seeking to promote flourishing, Seligman's (2011) WBT, has not been fully tested in a wide range of populations such that a true understanding of flourishing can be established. The current manuscript longitudinally tests key elements (e.g., the structure, reliability over time) of WBT. Then, this manuscript tests the ability of WBT to predict flourishing (e.g., high levels of well-being as indicated by better physical health and domain specific success) in two previously untested populations.

1.1 Well-Being Theory

Although there are many potential indicators of well-being, Seligman selected for WBT five aspects that people around the world pursue because they find them naturally rewarding (i.e., intrinsically motivating): positive emotions, engagement, relationships,

meaning, and achievement (a.k.a. PERMA). These five dimensions distinguish WBT from some other multidimensional theories (e.g., Diener et al. 2010; Ryan and Deci 2001; Ryff 1989) because it includes both eudemonic (i.e., living a purposeful life) and hedonic (i.e., living a life rich in joy and pleasure) components and includes unique components (e.g., engagement, achievement) that are less commonly included in other well-being theories. In contrast, other well-being theories include components, like vitality or resilience, which are products of people's active pursuits of well-being. In WBT, each of the individual PERMA dimensions work in concert to give rise to a higher-order construct of well-being that should then predict flourishing of groups, organizations, nations, or even the world (Forgeard et al. 2011; Seligman; 2011). For example, in one of the only direct tests of WBT, Kern et al. (2014) found significant positive associations between each PERMA indicators and physical health/vitality, life satisfaction, job satisfaction, and organization commitment in a sample of school employees. Thus, these five components have been individually linked to flourishing outcomes, but more research on PERMA is needed to determine whether the theorized higher-order structure holds (i.e., whether the five indicators form a broader well-being construct) and whether PERMA is linked to flourishing across multiple samples.

Positive emotions are the good things that we feel, such as happiness, hope, and joy (Cohn and Fredrickson 2009; Fredrickson 2001; Seligman 2011). Experiencing positive emotions is a primary goal of individuals around the world (Diener 2000). And, research suggests positive emotions are a key indicator of well-being (Coffey et al. 2014; Cohn and Fredrickson 2009; Lyubomirsky et al. 2005); they are positively related to life satisfaction, resilience, mindfulness, social rewards, work outcomes, and physical health (for a review, see Cohn and Fredrickson 2009).

A second important indicator of well-being is *engagement*—the act of becoming highly absorbed, interested, or focused in life activities (Csikszentmihalyi 1988). Flow, the prototypical form of engagement, occurs when participating in any absorbing activity that is highly challenging and requires a high level of skill; these attributes combine to lead to a loss of self-consciousness (Csikszentmihalyi 1990, 2008). Engagement is positively related to other indicators of well-being, including life satisfaction (Han 1988), satisfaction in work and leisure (Lefevre 1988), and increases in positive affect after the engagement experience (e.g., Rogatko 2009). Engagement increases academic commitment and achievement in high school students (Carli et al. 1988) and end-of-semester academic performance in college students (Engeser et al. 2005).

Feeling valued by others and having close, mutually satisfying *relationships* is another key indicator of well-being (Ryan and Deci 2000; Seligman 2011). Humans have a biologically anchored, innate propensity to form affectionate relationships throughout their lives, beginning at a very early age (Bowlby 1982). Likewise, some argue that close relationships represent a fundamental human need (e.g., Peterson 2006; Ryan and Deci 2000). Research has shown that relationships with friends are positively associated with self-esteem and that perceived increases in friendship quality are related to increases in well-being (Bagwell et al. 2005). Furthermore, college students who socialize more frequently and who have stronger romantic and social relationships tend to be happier than students without these relationships (Diener and Seligman 2002). Moreover, a study that included a sample representative of three-fourths of the world's population across 55 nations found that a good relationship was the only common predictor of happiness (Diener and Oishi 2000).

A fourth indicator of well-being is *meaning*—having a sense of purpose derived from something viewed as larger than the self (Seligman 2011; Steger et al. 2009). Independent

of its source (e.g., religion, relationships), people pursue meaning because it makes life worth living and gives them a sense of fulfillment (Chalofsky and Krishna 2009; Seligman 2011). Meaning is associated with other indicators of well-being throughout the adult lifespan (Steger et al. 2009), and relates to greater life satisfaction (Chamberlain and Zika 1988), higher rates of happiness, and fewer psychological problems (Debats et al. 1993).

Striving for *achievement* (or accomplishment), the fifth indicator of well-being, is described by Seligman (2011) as a persistent or determined drive to master or accomplish something for one's own sake. Thus, achievement as an indicator of well-being should be operationalized by examining people's desire to accomplish something (e.g., persevering attitude) rather than by examining actual accomplishments. Operationalizing achievement as holding a persevering attitude thus best aligns with WBT and extant research on PERMA (Kern et al. 2014; Seligman 2011). It also provides a subjective measure of achievement that is generalizable across contexts and ages and ensures that striving for achievement is a current and ongoing process. Conversely, actual achievements do not always lead to increased well-being (e.g., Grant and Dweck 2003), and thus there is not broad agreement on what kinds of accomplishments are valued. Perseverance predicts educational attainment and other accomplishments beyond that of IQ and conscientiousness (Duckworth et al. 2007). Furthermore, perseverance is positively related to grade point average (GPA), life satisfaction, and participation in extracurricular activities (Duckworth et al. 2007; Duckworth and Quinn 2009; Peterson et al. 2007).

1.2 The Current Project

Although WBT (Seligman 2011) provides a unique and exciting model for predicting flourishing through intrinsically motivated, eudaimonic and hedonic factors, key elements of the theory have yet to be tested. The structure as proposed by Seligman such that the five PERMA indicators should load onto a higher-order factor is untested. Furthermore, the long-term stability of PERMA has not been examined. Ultimately, the ability of the higher-order PERMA factor to predict elements of flourishing (e.g., physical health, personal success) over time needs to be tested. Before considering WBT worthy of further scientific study and before applying its tenets, we need to empirically test its veracity. The current paper offers a test of the higher-order structure of PERMA and its links to flourishing in two distinct populations. Our study is important for garnering evidence of WBT that will enable empirical comparisons to other multidimensional theories as well as further consideration of viability for interventions to promote flourishing.

Specifically, we evaluated the structure and stability of WBT within a college setting over three academic years (i.e., Study 1) and with a larger, more diverse community sample (i.e., Study 2). To test WBT in a college setting, we extracted items from an existing longitudinal data set to assess four of the five dimensions of WBT—positive emotions, engagement, relationships, and achievement (meaning was unavailable in this secondary dataset)—to examine the construct validity and temporal stability of the higher-order factor of well-being as proposed by WBT. Then, as a true test of PERMA's effectiveness in predicting flourishing, the higher-order factor was used to predict both concurrent and prospective physical health and objective indicators of college success (i.e., job interviews and GPAs). Study 2 utilized an online survey to (1) to further validate the PERMA measures used in Study 1, (2) add meaning to make a complete measure of PERMA, and (3) evaluate the generalizability of PERMA to a broader context by using it to predict flourishing. Taken together, these two studies provide empirical support for

WBT's core idea: there are multiple paths by which individuals pursue well-being and these pursuits combine to contribute to flourishing.

2 Study 1

The current study sought to examine the validity and predictive ability of Seligman's proposed WBT. Using existing longitudinal data collected at a small residential college, we examined four key questions: (1) Can a multidimensional model of PERMA with separate indicators that combine to form a higher-order well-being construct be adequately measured? We predicted that a second-order model of PERMA would demonstrate good model fit in a college setting (Hypothesis 1). (2) Will PERMA remain stable over time? We predicted that sophomore year PERMA would demonstrate increasing rank-order stability with junior and senior year PERMA (Hypothesis 2). (3) Will PERMA demonstrate construct validity through associations with other indicators of well-being? We predicted sophomore year PERMA would demonstrate positive prospective associations with other established indicators of well-being, including satisfaction with life and satisfaction with the college at the end of senior year, as well as changes in vitality and psychological distress over time (Hypothesis 3). (4) Will PERMA predict other concurrent and prospective indicators of flourishing in college? As we have documented the individual indicators of PERMA are often associated with physical health and other measures of domain specific success that are indicative of flourishing. Therefore, we predicted that PERMA would prospectively predict physical health over 2 years (Hypothesis 4a) and college success as indicated by GPA and number of interviews received at time of graduation (Hypothesis 4b).

2.1 Method

2.1.1 Participants

Data come from a 4-year, six-wave longitudinal study of college students with the goal of understanding community connection in relation to college experiences. In 2006, all incoming freshman of a small, undergraduate, private college focused on engineering, science, and math were invited to participate in the study; 119 participated at Wave I as part of their college orientation. Later, all non-participants were invited to enroll in the study at the end of their first semester (Wave II); 30 additional students participated, bringing total enrollment to 149 (82.3 % of the freshman class). Additional assessments occurred after the first year in college (Wave III: May 2007), and annually each subsequent year at the end of spring semester (Wave IV: May 2008, Wave V: May 2009, Wave VI: May 2010). Over the course of the study, 15 students were removed from the study (13 withdrew from the institution, 2 became privy to study hypotheses). Students were compensated $10 for Waves I–V and $20 for Wave VI. Data were collected via paper-and-pencil survey[1] in Wave I and computer-based self-report surveys in later waves.

At Wave II, participants ranged in age from 17 to 20 years ($M = 18.10$ years, $SD = .40$), and, consistent with the demographics of the college, most participants were male (69 %). Sample ethnicity was 53.8 % White, 17.3 % Asian, 9.6 % Hispanic, 1.3 %

[1] We had planned computer-based data collection for Wave I, but a technical difficulty made paper and pencil implementation necessary. Importantly, Wave I data are not at the centerpiece of this report.

Black, 1.3 % Native American, and 16.7 % not reported. Additionally, 2.6 % of the sample identified as non U.S. citizens. Only Waves IV-VI included the indicators of interest. Participation rates[2] for Waves IV, V, and VI were 89.4 % (126/141), 72.7 % (96/132), and 81.5 (106/130), respectively: hereafter, these waves are referred to as sophomore, junior, and senior year, respectively.

2.1.2 Well-Being Theory Indicators

Using Seligman's (2011) definitions of the five WBT indicators, we identified 40 theoretically relevant items within the existing data set that were included in sophomore, junior, and senior year surveys that reflected positive affect, engagement, relationships, meaning in life, and achievement. Most items came from empirically validated scales (e.g., Positive and Negative Affect Scale; Watson et al. 1988). Three researchers unfamiliar with the hypotheses classified these 40 items into five bins reflecting the WBT indicators and according to theoretical definitions provided by Seligman; average inter-rater reliability coefficient (Cohen's kappa) was .73. We used items that focused on periods between 1 week and one semester of time to reduce the influence of momentary emotion on our results. Given the college sample, we used relationship items that focused on their connectedness with other people on campus. Items that measured meaning could not be statistically identified; due to inadequate measures, this dimension was necessarily excluded. PERMA factors were tested through confirmatory factor analysis. Because of limited sample size, constructs were restricted to three to five items by focusing on items with higher inter-rater agreement to limit the number of model parameters while still representing the breadth of each latent construct. Items that did not significantly load were removed. Three items each were identified for positive affect, engagement, relationships, and achievement (all selected items appear in Appendix 1). At sophomore year, alphas were generally considered acceptable for a short scale (e.g., Briggs and Cheek 1986; Martin et al. 2002): .79 for positive affect, .66 for engagement, .52 for relationships, and .62 for achievement.

2.1.3 Measures to Assess Construct Validity

Composites were created for each of the following measures to reduce model complexity.

2.1.3.1 Vitality Vitality was measured using the 6-item Subjective Vitality Scale developed by Ryan and Frederick (1997) and modified by Bostic et al. (2000). Both manuscripts identify good internal consistency (α range from .80 to .89 over multiple studies). An example item is "Sometimes I feel so alive I just want to burst," and response options ranged from 1 (*not at all true*) to 7 (*very true*). Vitality was assessed at sophomore ($M = 4.22$, $SD = 1.27$; $\alpha = .91$) and senior years ($M = 4.39$, $SD = 1.19$; $\alpha = .87$).

2.1.3.2 Psychological distress We averaged four items from the Center for Epidemiological Studies Depression Scale (Radloff 1977) as a measure of psychological distress (i.e., feeling depressed, sad, lonely, or fearful). We could not use the full measure because items were used for PERMA indicators, so we focused on a limited number focused on

[2] Participation rates were computed as the total number of participants divided by the number of eligible participants; the number of eligible participants varied across waves due to leaves of absence, transfers away from the institution, and withdrawals from the study.

high-arousal negative affect. Response options ranged from 1 (*rarely or none of the time, <1 day*) to 4 (*most or all of the time, 5–7 days*). Data came from sophomore ($M = 1.75$, $SD = .71$; $\alpha = .83$) and senior year ($M = 1.68$, $SD = .61$; $\alpha = .75$).

2.1.3.3 Life satisfaction The Satisfaction with Life Scale, a 5-item self-report measure of an individual's global judgment of contentment in life, was used to measure life satisfaction (Diener et al. 1985). The authors reported good test–retest reliability ($r = .82$) and internal consistency ($\alpha = .87$). An example item is "The conditions of my life are excellent," and responses ranged from 1 (*strongly disagree*) to 7 (*strongly agree*). Satisfaction with life was assessed at senior year only ($M = 4.66$, $SD = 1.33$; $\alpha = .85$).

2.1.3.4 Satisfaction with college Satisfaction with college was measured with a single item at senior year ($M = 5.75$, $SD = 1.61$), "How satisfied are you with your decision to attend [this college]?" Responses ranged from 1 (*not at all satisfied*) to 7 (*very satisfied*). Past research supports single item indicators of satisfaction within a specific domain (Sandvik et al. 1993), and in this case more elaborate measures of college satisfaction were unavailable.

2.1.4 Measures to Assess Flourishing

2.1.4.1 Physical Health Physical health was assessed first by participant reported number of visits to see a health care provider for physical illness during spring semester at sophomore ($M = 0.52$, $SD = 1.32$) and senior year ($M = 0.70$, $SD = 2.18$). Second, for senior year ($M = 12.63$, $SD = 7.04$), participants completed the 57-item Pennebaker Inventory of Limbic Languidness (PILL Pennebaker 1982). Respondents indicated how frequently they experienced each symptom on a 5-point scale (1 = *have never or almost never experienced the symptom*, 2 = *<3 or 4 times per year*, 3 = *every month or so*, 4 = *every week or so*, or 5 = *more than once every week*). PILL scores were calculated for participants by summing the number of symptoms that participants reported experiencing every month or more (i.e., each response of 3, 4, or 5 added 1 to their total PILL score that started at zero).Thus, higher PILL scores indicated more symptoms ($\alpha = .89$).

2.1.4.2 Grade Point Average Students consented to have their GPA for each semester released by the campus registrar, providing an objective measure of domain success. Spring semester GPA was used for sophomore ($M = 3.10$, $SD = .51$) and senior year ($M = 3.32$, $SD = .40$).

2.1.4.3 Post-graduate Opportunities Participants reported the number of post-graduate interviews they had been offered for jobs, internships, and graduate school admission at the end of senior year ($M = 2.70$, $SD = 2.40$). These three items were summed to indicate post-graduate opportunities.

2.2 Results

2.2.1 Analytic Plan

Structural equation modeling (SEM) using M*plus* version 7.0 handled missing data using Full Information Maximum Likelihood (FIML) estimation. FIML, a model-based

procedure for reducing bias to parameter estimates due to missing data, maximally uses partial data such that all available response are used to estimate parameters (Enders 2010; Graham 2012). FIML yields unbiased parameter estimates and correct standard errors and likelihood ratios, often quite comparable to multiple imputation methods, particularly when missing data meet the assumption of missing at random or missing completely at random (MCAR; Graham and Coffman 2012). Little's MCAR test indicated a non-significant p value ($p = .29$), suggesting that MCAR is a plausible assumption for our missing data. In our models, there were between 11 and 19 missing data patterns and covariance coverage did not drop below .59 across parameter estimates, but in most coverage was between .72 and .97.

Our moderate sample size (maximum N for analyses was 138) has been deemed sufficiently large for modeling longitudinal data (Bentler 2007; Liu et al. 2012). Given SEM recommendations to not exceed a 10:1 ratio of 10 cases for each parameter estimated, we created composite indicators of each well-being indicator to reduce the number of parameters estimated overall and thus to avoid model over-identification (for more information, see Byrne 2012). Power calculations for the models tested in Study 1 ranged from .56 to .91 (with the majority above .80; Preacher and Coffman 2006). The majority of models used maximum likelihood (ML) estimation, and ML estimation with robust standard errors (MLR) was used when examining dependent variables with non-normal distributions (for a more detailed explanation of MLR, see King and Roberts 2014).

2.2.2 WBT Factor Structure

Correlations among composite indicators of well-being are shown in Table 1. A second-order factor of WBT was established using three indicators for each first-order latent construct: positive affect, engagement, relationships, and achievement (see Fig. 1). The final fit for sophomore year model was good [χ^2 (50) = 74.45, $p = .01$; CFI = .95; TLI = .93; RMSEA = .06] as suggested by Byrne (2012). Given the good fit, no other modifications were made. All component loadings were significant ($p < .001$). A follow up test was conducted where all 12 indicator items were loaded onto one first-order factor which yielded a marginal fit [χ^2 (54) = 103.37, $p < .001$; CFI = .89; TLI = .86; RMSEA = .09) that was significantly worse than the second-order factor; accordingly, all analyses from here will use the second-order WBT factor. Thus, in support of Hypothesis 1, we were able to establish empirical support for a second-order factor with four of the five theorized dimensions.

2.2.2.1 Longitudinal stability A longitudinal factor model was used to test the rank-order stability of PERMA across sophomore, junior, and senior years (see Fig. 2). Indicators were standardized before making composites. Error variances of composites were correlated with those of their respective composites across years. Model fit was good, $\chi^2(43) = 67.34$, $p = .01$; *CFI* = .97; *TLI* = .95; *RMSEA* = .06. No other modifications were made. As expected, sophomore year PERMA predicted junior year PERMA ($\beta = .73$, $p < .001$), which predicated senior year PERMA ($\beta = .90$, $p < .001$). The model predicted 54 % of the variance of junior year PERMA and 81 % of the variance at senior year. Tests of measurement invariance indicated equivalence of factor loadings ($\Delta\chi^2 = 4.20$, $\Delta df = 6$, $p =$ n.s.) and intercepts ($\Delta\chi^2 = 4.97$, $\Delta df = 14$, $p =$ n.s.) over time; thus, strong measurement invariance was established. In support of Hypothesis 2, the PERMA structure was increasingly stable across three academic years.

Table 1 Correlations among Study 1 sophomore and senior year composite variables

	1	2	3	4	5	6	7	8	9	10	11
(1) PE	**.64*** **	.64***	.32***	.68***		−.58***	−.54**			.11	−.21*
(2) Engagement	.59**	**.55*** **	.17	.60***		−.42***	.58***			.25**	−.20*
(3) Relationships	.27**	.14	**.55*** **	.22*		−.02	.37***			−.07	.08
(4) Achievement	.58**	.48**	.17	**.57*** **		−.55***	.54***			.26**	−.11
(5) Interviews	.16	.03	.04	.17	–						
(6) Psych Distress	−.52***	−.48***	−.07	−.32***	−.02	**.46*** **	−.29**			−.20*	.32***
(7) Vitality	.74***	.71***	.24*	.54***	.05	−.56***	**.63*** **			.22*	−.13
(8) SWL	.54***	.35***	.10	.50***	.00	−.42***	.59***	–			
(9) SWC	.36***	.41***	.26**	.21*	.00	−.30***	.34***	.34**	–		
(10) GPA	.24*	.23*	.04	.28**	.30**	−.20	.23*	.33***	.21*	**.94*** **	−.12
(11) Health	.17	.08	.04	.04	.14	−.10	.12	.04	.06	−.12	**.18**

N's range from 93 to 123 due to occasional missing data. Sophomore year correlations are represented in the top table and senior year correlations are on the bottom of the table. Bolded values between the top and bottom represent correlations between sophomore and senior year of that variable

PE positive emotions, *Interview* interviews for post-graduate opportunities, *Psych distress* psychological distress, *SWL* satisfaction with life, *SWC* satisfaction with college, *GPA* grade point average

*$p < .05$. ** $p < .01$. *** $p < .001$

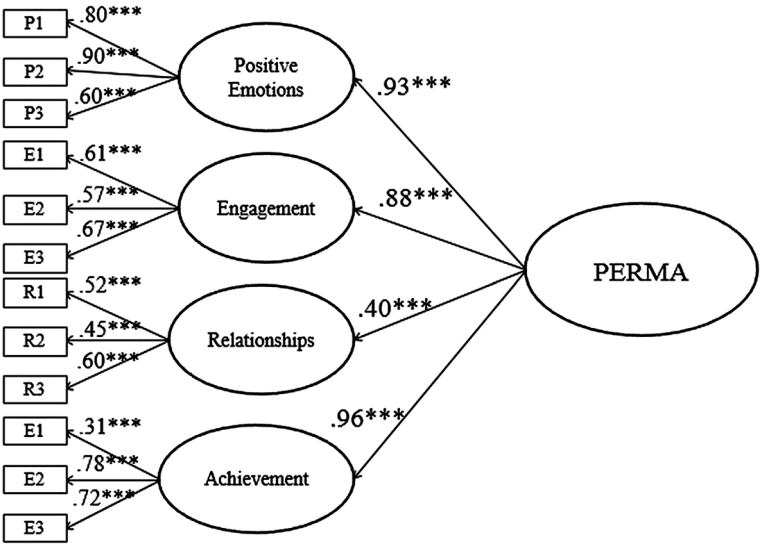

Fig. 1 Sophomore year second-order PERMA model ($N = 127$). *CFI* = .95; *TLI* = .93; *RMSEA* = .06; $\chi^2(50) = 74.45$, $p < .01$. Standardized factor loadings are displayed. *P* positive emotions, *E* engagement, *R* relationships, *A* achievement, *WBT* well-being theory. ***$p < .001$

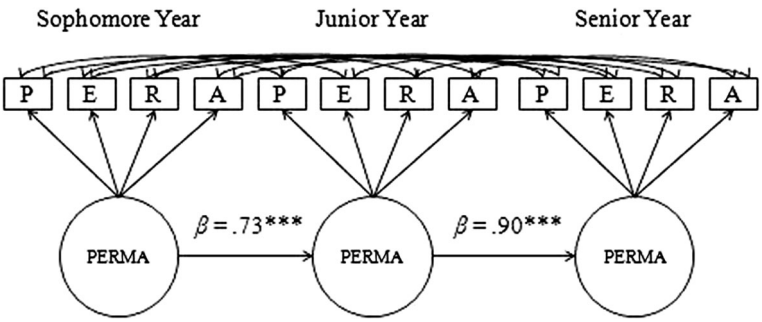

Fig. 2 Stability of WBT over three years ($N = 133$). *CFI* = .97; *TLI* = .95; *RMSEA* = .06; $\chi^2(43) = 67.34$, $p = .01$. Junior year $R^2 = .54$. Senior year $R^2 = .81$. *P* positive emotions, *E* engagement, *R* relationships, *A* achievement. ***$p < .001$

2.2.3 Construct Validity

2.2.3.1 Concurrent and Prospective Validity For analyses testing concurrent and prospective construct validity, we used sophomore year PERMA (see Fig. 1) to predict other indicators of well-being (i.e., vitality, psychological distress, satisfaction with life, and satisfaction with college) at the end of sophomore and senior year (see Fig. 3)[3]. The model fit was fair, χ^2 (123) = 221.16, $p < .001$; *CFI* = .88; *TLI* = .85; *RMSEA* = .08. As expected, sophomore year PERMA associated with sophomore vitality ($\beta = .63$, $p < .001$)

[3] For longitudinal analyses, we focused only on outcomes from senior year because it provided more outcomes of interest than junior year. In separate models, between sophomore and junior year, results were generally comparable to the ones reported in the manuscript, so we focused on the longer time period.

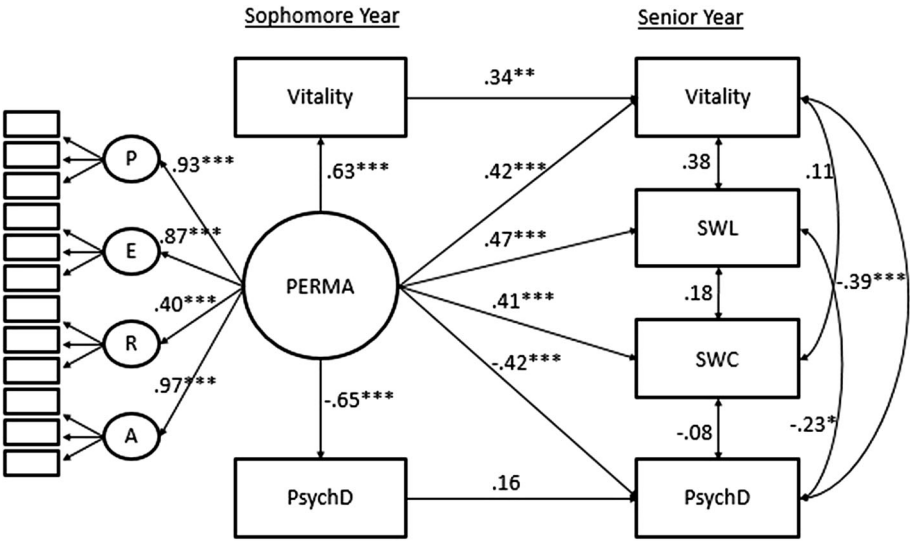

Fig. 3 WBT concurrent and prospective validity model ($N = 127$). Model fit was fair: $CFI = .88$; $TLI = .85$; $RMSEA = .08$; $\chi^2(123) = 221.16$, $p < .001$. See Fig. 1 for a complete representation of PERMA model. *P* positive emotions, *E* engagement, *R* relationships, *A* achievement, *PsychD* psychological distress, *SWC* satisfaction with college, *SWL* satisfaction with life. *$p < .05$; **$p < .01$; ***$p < .001$

and senior vitality ($\beta = .42$, $p < .001$). Vitality was fairly stable across the 2 years ($\beta = .35$, $p < .001$). Thus, PERMA predicted relative changes in vitality over 2 years after controlling for earlier vitality. The validity model accounted for 40 % of sophomore year variance and 48 % of senior year variance in vitality. Similarly, PERMA predicted sophomore year ($\beta = -.68$, $p < .001$) and senior year psychological distress ($\beta = -.44$, $p < .01$). Surprisingly, concurrent psychological distress did not associate with prospective psychological distress. PERMA accounted for 46 % of sophomore year variance in psychological distress and 30 % of senior year variance. As expected, PERMA positively related to prospective satisfaction with life and satisfaction with college: Sophomore year PERMA accounted for 22 % of variance of satisfaction with life ($\beta = .47$, $p < .001$) and 17 % of variance of satisfaction with college ($\beta = .41$, $p < .001$) in senior year. Thus, consistent with Hypothesis 3, sophomore year PERMA demonstrated construct validity concurrently and prospectively.

2.2.3.2 Indicators of Flourishing We used two indicators of flourishing—physical health and college success. We again used sophomore year PERMA to predict sophomore and senior year physical health appointments and senior year physical health symptoms (see Table 2). ML estimation with robust standard errors (MLR) was used to handle non-normality of physical health appointments. The model for physical health outcomes had good fit, $\chi^2(92) = 121.92$, $p < .01$; $CFI = 1.00$; $TLI > 1.00$ $RMSEA = .06$, so no modifications were made. PERMA predicted sophomore physical health appointments ($B = -1.12$, $p < .01$), but neither sophomore physical health appointments nor PERMA significantly predicted senior year physical health appointments. Sophomore PERMA significantly predicted senior year physical health symptoms ($B = -3.26$, $p < .05$). Thus, consistent with Hypothesis 4a, findings suggest that sophomore year PERMA predicted concurrent and prospective health.

Table 2 Sophomore year WBT predicting prospective physical health and college success

Pathways	Parameters	
Health model ($N = 130$)	*B*	*SE*
PERMA → sophomore PH appt	−1.21**	.38
PERMA → senior PH appt	−1.78	1.32
PERMA → senior pills	−3.26*	1.44
Sophomore PH appt → senior PH appt	−.33	.28

College success model ($N = 138$)	β	*SE*	R^2
PERMA → sophomore GPA	.26**	.09	.07
PERMA → senior GPA	−.03	.09	.00
PERMA → post-graduate opportunities	.27*	.11	.07
Sophomore GPA → senior GPA	.61***	.07	.36

Model uses second-order PERMA structure (see Fig. 1) to predict outcomes variables. Model fit for was adequate for PH model (χ^2 [92] = 121.92, $p < .01$; *CFI* = 1.00; *TLI* = 1.69; *RMSEA* = .06) and for the college success model [χ^2 (84) = 122.86, $p < .01$; *CFI* = .93; *TLI* = .91; *RMSEA* = .06]. Unstandardized estimates are reported and R^2 are not, due to using the MLR estimator to deal with non-normality of physical health (PH) appointments

PH physical health, *PILLS* pennebaker inventory of limbic languidness, *GPA* grade point average for spring term, *Appt* appointments, *SE* standard error

* $p < .05$. ** $p < .01$. *** $p < .001$

For analyses predicting objective flourishing outcomes for the college success domain, we used sophomore year PERMA to predict concurrent and prospective GPA and prospective post-graduate opportunities (see Table 2). This model had adequate fit, $\chi^2(84) = 122.86$, $p < .01$; *CFI* = .93; *TLI* = .91; *RMSEA* = .06. No model modifications were made. As expected, sophomore year PERMA predicted sophomore year spring semester GPA ($\beta = .26$, $p < .01$), but, unexpectedly, did not predict senior year spring GPA.[4] Sophomore year GPA significantly predicted senior year GPA ($\beta = .61$, $p < .001$). The model accounted for 7 % of variance in sophomore year GPA and 36 % of variance in senior year GPA. In addition, sophomore year PERMA predicated more interviews for post-graduate opportunities at senior year ($\beta = .27$, $p < .05$); this model accounted for 7 % of variance in post-graduate interviews. Thus, consistent with Hypothesis 4b, these findings suggest that sophomore year PERMA predicted concurrent GPA and prospective post-graduate interviews.

2.3 Discussion

Our longitudinal examination supports WBT's central tenet that (four of the five) different dimensions of well-being can be appropriately considered aspects of a single multidimensional structure. Furthermore, the PERMA model showed temporal rank-order stability, and, as expected, PERMA associated with other indicators of well-being, both

[4] GPA was not significantly related to sophomore year PERMA at any of the subsequent semesters, therefore, we only report on spring GPAs in sophomore and senior years for consistency with the temporality of the other dependent variables.

concurrently and prospectively. Importantly, PERMA predicted flourishing by linking to better physical health and college success 2 years later.

Yet, this study utilized secondary data, the available items were less than perfect assessments of the constructs, and the modest sample size limited the complexity of the testable models. Additionally, one construct from WBT (meaning) was not measurable due to the total absence of appropriate proxy items. Moreover, because the sample was comprised of students at an engineering, science, and mathematics college, it is possible that unique features of the sample limit the generalizability of these findings (even to other college students). Study 2 addressed these limitations.

3 Study 2

Study 2 examined three additional questions using data collected from a much larger and more diverse sample: (1) Do each of the Study 1 PERMA indicators show convergent and discriminant validity with established indicators of PERMA components? We predicted that Study 1 measures would demonstrate validity with their corresponding indicators (Hypothesis 5). (2) Will the PERMA model generalize to a broader population? We predicted that Study 1 PERMA measurement model would fit a broader community sample (Hypothesis 6). (3) Can we improve upon our PERMA model by adding meaning to the model from Study 1? We predicted that meaning would fit within the model (Hypothesis 7). Finally, we expected this final model would predict similar cross-sectional well-being and flourishing outcomes to our Study 1 PERMA model (Hypothesis 8).

3.1 Method

3.1.1 Participants

In 2013, 831 participants were recruited online for a cross-sectional study about emotions and well-being using an online survey platform. As we sought a more heterogeneous sample with a broad range of well-being that would also utilize a computer-based reporting system (similar to Study 1), we recruited participants through Amazon's Mechanical Turk (Mturk). Although not representative of the general population, Mturk offers a more diverse and older sample than college students alone, and this method is considered as reliable as traditional data collection methods (Buhrmester et al. 2011; Paolacci and Chandler 2014). Mturk populations are usually more educated, more liberal, less employed, and less religious than the general population. Furthermore, Mturk samples are honest, diligent, and reliable reporters for psychometric tests that are motivated for intrinsic and extrinsic purposes (for a review, see Paolacci and Chandler 2014). Participants were required to be located in the United States of America. Individuals received $5[5] for participating. A number of different survey design principles (e.g., reversing items, changing response options, verifiable answers) were utilized for quality assurance based on prior recommendations (e.g., Mason and Suri 2012; Paolacci and Chandler 2014). Data

[5] $5 is a much higher compensation rate than is common for a 30 min survey on Mturk which resulted in us receiving many comments from participants thanking us for paying a fair rate. Although, research suggests Mturk participants are usually vested, the comments led us to believe they were more vested given the higher compensation.

were examined for any abnormal patterns of responses and outliers or for large amounts of missing data. In total, seven cases were removed.[6]

Participants (56 % female) ranged in age from 18 to 75 years ($M = 35.59$ years, $SD = 11.81$); 76.2 % were White/Caucasian, 9.6 % African American, 5.4 % Asian, 3.1 % Hispanic, 1.3 % Native American, 2.9 % other, and 1.4 % unreported. A large portion of the sample (45.8 %) had earned a 4-year degree, 11.1 % had earned a 2-year degree, 31.5 % had some college, 11.0 % finished high school or earned a GED, and .6 % had not finished high school. Over a third of the sample was married (39.0 %), 36.3 % were single, 14.1 % were living with their partner, 8.6 % were divorced, 1.2 % were separated, and .9 % were widowed. Thus, this population was older and more diverse than the Study 1 population.

As with Study 1, SEM using M*plus* version 7.0 (Muthén and Muthén 1998–2012) handled missing data using FIML estimation. Maximum N for analyses was 824.

3.2 Measures

3.2.1 Well-Being Theory Indicators

As a reminder, Study 1 used items from a secondary dataset resulting in 3 items for each of the four first-order factors (i.e., positive emotions, engagement, relationships, and achievement) that comprised the second-order PERMA factor. Identical items were administered in Study 2 for positive affect, engagement, and achievement. The relationship items required modifications to make them applicable for this broader sample. Two of the three relationship items about the participants' relationship with their college community and dorm community were replaced by similarly structured questions asking about their relationship with their general friends and their community; the third relationship question remained the same. See Appendix 1 for details. Alphas were .78 for positive affect, .80 for engagement, .69 for relationships, and .71 for achievement.

3.2.2 Measures to Assess Convergent and Discriminant Validity

Study 1 was limited in relying on secondary data to create proxy measures of the PERMA constructs. Thus, a goal of Study 2 was to test whether these proxy measures mapped onto established "gold standard" measures of PERMA constructs; that is, convergent validity of the Study 1 items was examined in Study 2. Study 2 also tested discriminant validity of the Study 1 items by administering PERMA measures alongside established measures of constructs theoretically unrelated to PERMA.

3.2.2.1 Positive and Negative Affect Positive and negative affect were measured using the Affect Adjective Checklist (AAC; Diener and Emmons 1985). The authors reported good internal consistency for positive ($\alpha = .89$) and negative ($\alpha = .84$) affect. The AAC has been previously validated and used to validate many other measures of positive and negative affect (e.g., Watson et al. 1988). Respondents indicated how often they felt positive affect (e.g., joyful) and negative affect (e.g., frustrated) over the past week on a 7-point scale from 1 (*not at all*) to 7 (*extremely*). Positive ($M = 4.44$, $SD = 1.42$; $\alpha = .94$) and negative affect ($M = 3.51$, $SD = 1.65$; $\alpha = .90$) were used to test convergent validity

[6] No attention checks were used meaning that unidentified deviant responders could add error to the estimates.

and discriminant validity (using negative affect) of the Study 1 items used to tap positive affect.

3.2.2.2 Engagement Participants' flow experience (Csikszentmihalyi 1988, 1990) was measured using an established 5-item scale designed for measuring engagement over the last week (e.g. Layous et al. 2013; Nelson et al. 2014). Past studies have provided evidence of validity and good internal consistency (α ranging from .77 to .87; Layous et al. 2013; Nelson et al. 2014). Participants responded to these 5 items (e.g., "I felt absorbed in what I was doing") about experiences in the past week on a 7-point scale, 1 (*not at all*) to 7 (*very much*). Flow ($M = 4.20$, $SD = 1.04$; $\alpha = .66$) was used to test convergent validity of the Study 1 engagement measure.

3.2.2.3 Relationships The Relatedness subscale of the Basic Need Satisfaction Scale (BNSS; Gagné 2003) was used as a measure of relationships. Internal consistency was good (α range from .69 to .86) on this widely used and validated measure of how connected a person feels to others (e.g., Gagné 2003; Kashdan et al. 2006). Participants responded to 8 items on a 7-point scale from 1 (*not at all true*) to 7 (*very true*) about how true items like "People in my life care about me." Relatedness ($M = 6.17$, $SD = 1.07$; $\alpha = .85$) was used to test convergent validity of the Study 1 relationships measure.

3.2.2.4 Achievement The commonly used and previously validated Short Grit Scale was used to measure perseverance of effort, which is most conceptually similar to the component of achievement we assessed in Study 1 (Duckworth and Quinn 2009). The scale had 6 items (e.g., "I finish whatever I begin") and adequate internal consistency ($\alpha = .60$ to .78). Participants responded on a 5-point scale from 1 (*not at all like me*) to 5 (*very much like me*). Perseverance of effort ($M = 3.81$, $SD = .72$; $\alpha = .81$) was used to test convergent validity for our Study 1 achievement measure.

3.2.2.5 Meaning A measure of meaning was collected using the widely used 10-item Meaning in Life Questionnaire (MIQ; Steger et al. 2006), which includes two subscales: (1) The Presence of Meaning (e.g., "My life has a clear sense of purpose."), and (2) the Search for Meaning (e.g., "I am always looking to find my life's purpose."). Internal validity and consistency was good for the presence of meaning (αs ranged from low to high .80 s) and for the search for meaning (αs ranged from mid .80 s to low .90 s; Steger et al. 2006). The 7-point response options range from 1 (*absolutely untrue*) to 7 (*absolutely true*). Although, only the presence of meaning ($M = 4.59$, $SD = 1.60$; $\alpha = .95$) will be used for the PERMA model, search for meaning was also measured for discriminant validity testing ($M = 4.51$, $SD = 1.58$; $\alpha = .94$) because the search for meaning is distinct from the presence and usually related to feeling like life does not have a purpose; search for meaning is often unrelated or inversely related to positive affect, presence of meaning, and life satisfaction and positively related to negative affect and depression (e.g., Steger et al. 2006).

3.2.3 Study 1 Measures Used for Replication

3.2.3.1 Validity The same Study 1 measures of vitality ($M = 4.53$, $SD = 1.47$; $\alpha = .93$), psychological distress ($M = 1.69$, $SD = .78$; $\alpha = .87$), and life satisfaction ($M = 4.21$, $SD = 1.53$; $\alpha = .93$) were used to measure validity.

3.2.3.2 Flourishing Physical health symptoms ($M = 15.0$, $SD = 10.1$; $\alpha = .93$) were measured the same way as in Study 1.

3.3 Results

3.3.1 Validity Testing

Convergent and discriminant validity of the proxy measures derived from Study 1 was tested by correlating these composites with established measures of the well-being dimensions. Importantly, all Study 1 PERMA indicators were positively correlated with the gold standard measures of the same/similar construct at .54 ($p < .001$) or above (see Table 3). As such, the proxy measures employed in Study 1 evidenced acceptable convergent validity with established measures.

Discriminant validity of Study 1 proxy measures was evaluated by examining correlations with measures that have theoretically limited relationships to PERMA. None of our PERMA composites were significantly related to sex, ethnicity, or education ($r < .07$; $p > .05$). These findings help to establish the independence of our PERMA measures. In addition, prior work has demonstrated that positive affect inversely correlated with negative affect (e.g., Diener and Emmons 1985) just as it was in our data ($r = -.61$, $p < .001$). As expected, the other PERMA variables were also inversely related to negative affect (r's ranging from $-.45$ to $-.31$, $p < .001$). Furthermore, previous work has indicated that search for meaning has a nonsignificant or small association with our PERMA indicators of well-being (e.g., Steger et al. 2006), and likewise associations between PERMA and search for meaning were weak (r's ranging from $-.16$ to $-.08$, $p < .01$). In summary, results supported Hypothesis 5 and give evidence of construct validity for PERMA indicators.

3.3.2 WBT Factor Structure

As an expansion of Study 1, a second-order factor of WBT was established using the same latent constructs as in Study 1 (i.e., positive affect, engagement, relationships, achievement) for a sub-sample of participants under 24 years old ($n = 130$; Model 1). Model 1 fit was acceptable without any modifications [$CFI = .93$; $TLI = .91$; $RMSEA = .08$; χ^2 $(50) = 92.93$, $p < .001$]. Model 2, which utilized the full sample, demonstrated marginal fit; correlating the errors of two achievement items improved model fit to an acceptable level [$CFI = .93$; $TLI = .90$; $RMSEA = .09$; χ^2 $(49) = 359.85$, $p < .001$]. For both models, all factor loadings exceeded .50, $p < .001$. Therefore, Hypothesis 6 was supported by replicating our Study 1 model with a similarly aged, emerging adult sample (Model 1) and for a broader population (Model 2) by adding one modification (correlating errors of two achievement items).

3.3.3 Addition of Meaning to WBT Model

To offer a more complete examination of WBT, we tested the PERMA model after adding meaning (Model 3) to Model 2. Model fit was acceptable [$CFI = .94$; $TLI = .93$; $RMSEA = .08$; χ^2 $(114) = 684.30$, $p < .001$] without any modifications (see Fig. 4). Thus, the full structure of PERMA was verified. Hypothesis 7 was supported.

Table 3 Means, standard deviations, and correlations among Study 2 composite variables use for convergent and divergent validity

	Mean	*SD*	1	2	3	4	5	6	7	8	9	10
(1) PE	2.74	1.08										
(2) Engagement	3.95	0.73	.61***									
(3) Relationships	4.20	1.04	.47***	.36**								
(4) Achievement	4.44	1.43	.50***	.64***	.36***							
(5) Flow	3.51	1.65	.47***	.54***	.28***	.47***						
(6) AAC-PA	6.17	1.07	.88***	.63***	.48***	.50***	.48***					
(7) AAC-NA	5.11	1.39	−.62***	−.45***	−.35***	−.38***	−.31***	−.61***				
(8) Relatedness	3.04	0.89	.59***	.47***	.64***	.45***	.31***	.61***	−.51***			
(9) Perseverance	4.21	1.53	.46***	.57***	.38***	.72***	.40***	.47***	−.42***	.46***		
(10) Meaning	1.69	0.78	.58***	.51***	.47***	.55***	.40***	.56***	−.49***	.57***	.55***	
(11) Searching	14.97	10.11	−.16***	−.12***	−.08*	−.10**	.01	−.16***	.26***	−.15***	−.12**	−.33***

PE positive emotion, *AAC* affect adjective checklist, *PA* positive affect, *NA* negative affect, *Master* mastery goals, *Searching* searching for meaning, *PsychD* psychological distress

* $p < .05$. ** $p < .01$. *** $p < .001$

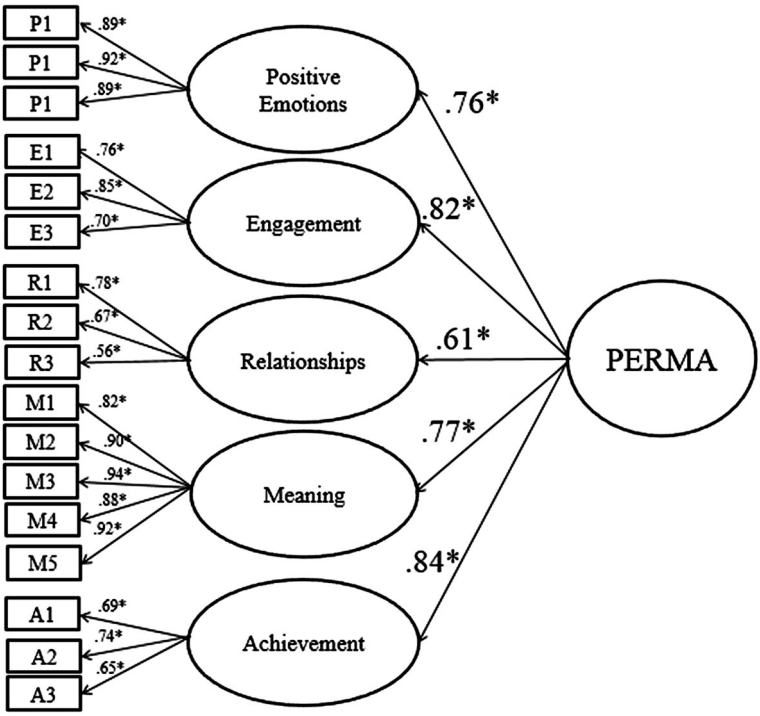

Fig. 4 Community sample ($N = 831$) second-order PERMA Model. χ^2 (114) = 684.30, $p < .001$; $CFI = .94$; $TLI = .93$; $RMSEA = .08$. Two of the achievement error terms were correlated (not shown). All standardized factor loadings are displayed. *P* positive emotions, *E* engagement, *R* relationships, *A* achievement. $^*p < .001$

3.3.4 WBT Associations

In Model 4, we used Model 3 to predict vitality, life satisfaction, and psychological distress. The model demonstrated acceptable fit [$CFI = .93$; $TLI = .91$; $RMSEA = .08$; χ^2 (162) = 1,015.66, $p < .001$]. The higher-order PERMA factor accounted for significant amounts of variance in vitality (85 %), life satisfaction (64 %), and psychological distress (52 %).

Similarly, in Model 5 (flourishing model), we used Model 3 to predict physical health symptoms. Model 5 demonstrated acceptable fit ($TLI = .92$; $RMSEA = .07$ χ^2 [130] = 728.21, $p < .001$; $CFI = .94$). PERMA accounted for 10 % of variance in physical health. Hypothesis 8 was supported in replicating links from PERMA to other well-being and flourishing outcomes found in Study 1.

3.4 Discussion

Study 2 extended the Study 1 findings by (1) demonstrating that the Study 1 items have reasonable convergent validity with the gold standard items; (2) showing that the structure of the PERMA model replicated in this older and more diverse sample; and (3) illustrating

that the full PERMA model was verified by adding a meaning indicator, which resulted in an acceptable-fit to the data. Finally, the concurrent associations between PERMA and Study 1 indicators of well-being and flourishing were replicated. As for flourishing, PERMA predicted 10 % of the variance in physical health.

4 General Discussion

The current research adds to the current empirical support for Seligman's (2011) WBT. Results from two studies with different samples and using cross-sectional and longitudinal data converge to provide strong evidence for the theorized multidimensional structure of well-being and its links to flourishing. Key findings suggest that well-being can be validly measured with the five dimensions of positive affect, engagement, relationships, meaning and achievement; WBT has substantial rank-order stability over multiple years; and PERMA is useful for predicting physical health and college success. Collectively, these findings contribute to understanding the theoretical and practical value of WBT at the organization level (e.g., a college) and in a more diverse, community sample.

4.1 Support for Well-Being Theory

Supporting the burgeoning theoretical views that well-being is best conceptualized multidimensionally (e.g., Huppert and So 2013; Ryan and Deci 2000; Ryff 1989; Seligman 2011), we provided empirical support for WBT. A higher-order PERMA model, with and without meaning, demonstrated adequate to good model fit. The more limited PERMA model was increasingly stable from sophomore to junior to senior year based on rank-order coefficients: in other words, individual differences in PERMA became more stable over time. Importantly, tests of measurement invariance revealed that WBT can be measured consistently across the college years. These tests are a significant step toward empirically and longitudinally testing a measurement model rooted in WBT (Seligman 2011), and we demonstrated a fairly robust multidimensional model of well-being.

Some convincing evidence of validity of our PERMA model emerged across the two studies. In both studies, PERMA was associated with indicators of well-being commonly used in hedonic studies of well-being such as life satisfaction (e.g., Diener 2000). PERMA was also associated with changes in vitality over 2 years. PERMA accounted for between 40 and 85 % of the variance in vitality in our studies. Moreover, PERMA was more strongly associated with prospective psychological distress than individuals' past reports of their psychological distress. This finding supports the idea that researchers, clinicians, and educators should seek to examine what is going well (i.e., WBT), as this way of measuring well-being is helpful for understanding both ends of the mental health spectrum.

Consistent with previous research examining individual differences in psychological well-being (e.g., Huppert and So 2013; Keyes 2007), higher levels of PERMA in college students and a more diverse community sample indicated better physical health. Indeed, PERMA concurrently predicted fewer physical health-related doctor visits and predicted fewer self-reported physical health symptoms concurrently and 2 years later. Sophomore year PERMA did not predict senior year health visits, but neither did sophomore year health visits. The lack of rank-order stability suggests high levels of variability in health

visits over time for this college population, which might explain why PERMA did not predict prospective health visits. Taken together, our results suggest that well-being, conceptualized and measured as PERMA, has direct relevance for physical health.

Importantly, well-being should be assessed in a way that relates to individuals' flourishing within their own environments—in Study 1, this environment was a rigorous academic setting. Evidence emerged in favor of immediate as well as downstream academic benefits of the limited measure of PERMA. Specifically, PERMA associated with participants' concurrent (but not prospective) GPA. Sophomore PERMA also prospectively predicted interviews for post-graduate opportunities at the end of senior year. As expected based on a multidimensional model of well-being, our results demonstrated that PERMA is useful for predicting objective, domain specific flourishing for college students. It is also possible, yet untested, that PERMA better equips college students with the psychological resources needed for life after college.

Notably, PERMA accounted for modest amounts (e.g., about 7 %) of GPA and post-graduate interviews. These findings are particularly striking as they represent more distal, domain specific measures. Furthermore, GPA was collected from university records (not self-reported) at a rigorous and competitive university where detecting even 7 % of the variance is difficult.

These two studies make important theoretical contributions to positive psychology by providing empirical support for the structure and longitudinal stability of a multidimensional conceptualization of well-being, as well as its empirical links to flourishing. WBT is unique in relation to other well-being theories because it combines intrinsically motivated indicators of eudaimonic and hedonic well-being and includes aspects of well-being that individuals can actively and agentically pursue.

4.2 Practical Applications of WBT

Assuming communities and colleges strive to provide a well-balanced and rich experience, officials and administrators could endeavor to enhance positive affect, relationships, engagement, meaning, and achievement by promoting a vast array of activities known to promote multiple aspects of well-being that people enjoy pursuing. Colleges could use WBT to track and report their own ability to promote well-being, which administrators could use to assess the university climate and identify students who might be in danger of experiencing psychological distress, poor health, or academic problems. Communities and other institutions that wish to offer programmatic support for well-being may benefit from understanding the benefits of WBT and how best to tailor activities to promote flourishing (e.g., Nelson and Lyubomirsky 2012). Thus, WBT may offer a simple way to study easy-to-understand indicators of well-being, and these indicators may be tangible enough to be potentially amenable to focused intervention efforts.

4.3 Limitations and Future Directions

Several limitations are worth noting. The sample size for Study 1 restricted our capacity to test more complex models; we had to keep our models more parsimonious to avoid model under-identification issues. Fortunately, given large effect sizes among variables, our

power was moderate to high in Study 1 despite the low sample size. Study 2 was cross-sectional, and therefore we could not fully replicate the lagged associations in Study 1. Both samples were limited in their generalizability and were not fully representative of their respective populations. Furthermore, all variables (except GPA) in both studies were self-reported when a more diverse range of collection methods would have been more optimal. Study 1 was limited because the necessary variables were only available at the end of sophomore, junior, and senior years. More frequent collection of certain variables (e.g., emotions) at different times throughout the year and including freshmen year would have yielded more detailed longitudinal information. Similarly, a lack of post-college data for Study 1 prohibits exploration on the relationship between PERMA during college and post-college flourishing.

Despite these limitations, these studies have a number of strengths. First, secondary data analysis of PERMA offered a conservative test of the model given that measures were not designed to test the theory. Study 2 addressed sample size, measurement, and sampling limitations of Study 1, and the model still showed adequate fit, stability, and validity through a rigorous variety of tests with the two samples. Furthermore, the model operated as hypothesized in both studies and associated with other indicators of well-being and predicted flourishing (e.g., health, and college success). While monolithic studies of WBT dimensions will remain important, our results provide further evidence that using a multidimensional model of well-being will continue to broaden our understanding of well-being and its implications for everyday life in which people pursue well-being in many ways.

Now that initial empirical support for WBT has been demonstrated with clear links to flourishing behaviors, researchers should continue to explore this multidimensional approach in basic and applied research. Future measurement work could incorporate item response theory to better determine the most precise and efficient measures for capturing PERMA. Theory-building work could compare WBT to other multi-dimensional models (e.g., Ryan and Deci 2000; Ryff 1989) to determine which are most effective and easily applied to communities. More research is needed to further elucidate the benefits of WBT for health and other domains over longer periods of time in a greater variety of contexts, such as the work place. Applied research is needed to identify the life experiences, programs, and practices that bolster multiple dimensions of well-being. Although broader application testing WBT needs to be done, our empirical support suggests that WBT may be poised to provide organizations, communities, and nations with easy to understand indicators of well-being that can be applied to individuals seeking to promote flourishing.

Acknowledgments Study 1 was supported by an internal grant from Harvey Mudd College and Study 2 was supported by an internal grant from Claremont Graduate University. Gratitude is extended to the Harvey Mudd Students and Mturk workers that participated in this study, and to, Katie Nelson, Mihaly Csikszentmihalyi, Jeanne Nakamura, Jessica Borelli, David Kyle Bond, Thomas Chann, Binghuang A. Wang, and Vicky Bouche for their assistance on this manuscript.

Appendix

See Table 4.

Table 4 Individual items for original first-order factors of well-being for both studies

Item	Response options	Well-being factor
I was happy	1 = *rarely or none of the time (<1 day)* to 4 = *most or all of the time (5–7 days)*	P
Feel in the last week: excited	1 = *very slightly or not at all* to 5 = *extremely*	P
I enjoyed life	1 = *rarely or none of the time (<1 day)* to 4 = *most or all of the time (5–7 days)*	P
Feel in the last week: attentive	1 = *very slightly or not at all* to 5 = *extremely*	E
Feel in the last week: interested	1 = *very slightly or not at all* to 5 = *extremely*	E
Feel in the last week: active	1 = *very slightly or not at all* to 5 = *extremely*	E
Select the picture that best describes your relationship with the college community[a]	7 progressively overlapping circles	R
Select the picture that best describes your relationship with the people in your dorm community[a]	7 progressively overlapping circles	R
Circle the picture that best describes your relationship with your friends, in general[b]	7 progressively overlapping circles	R
Circle the picture that best describes your relationship with your community[b]	7 progressively overlapping circles	R
I enjoy social gatherings just to be with people	1 = *not at all characteristic of me* to 5 = *very much characteristic of me*	R
Feel in the last week: Determined	1 = *very slightly or not at all* to 5 = *extremely*	A
If at first you don't succeed try, try again	1 = *strongly disagree* to 5 = *strongly agree*	A
I always try to persevere no matter what	1 = *strongly disagree* to 5 = *strongly agree*	A

[a] Is unique to Study 1

[b] Is unique to Study 2. *P* positive emotions, *E* engagement, *R* relationships, *A* achievement. In Study 1 meaning was excluded from the model because of a lack of adequate items measuring meaning. In Study 2 the Meaning in Life Questionnaire was used (Steger et al. 2006). Generally, all instructions asked participants to consider the last week

References

Bagwell, C., Bender, S., Andreassi, C., Kinoshita, T., Montarello, S., & Muller, J. (2005). Friendship quality and perceived relationship changes predict psychosocial adjustment in early adulthood. *Journal of Social and Personal Relationships, 22*, 235–254. doi:10.1177/0265407510380607.

Bentler, P. M. (2007). On tests and indices for evaluating structural models. *Personality and Individual Differences, 42*, 825–829. doi:10.1016/j.paid.2006.09.024.

Bostic, T., Rubio, D., & Hood, M. (2000). A validation of the subjective vitality scale using structural equation modeling. *Social Indicators Research, 52*, 313–324. doi:10.1023/A:1007136110218.

Bowlby, J. (1982). Attachment and loss: Retrospect and prospect. *American Journal of Orthopsychiatry, 52*(4), 664–678. doi:10.1111/j.1939-0025.1982.tb01456.x.

Briggs, S., & Cheek, J. (1986). The role of factor analysis in the development and evaluation of personality scales. *Journal of Personality, 54*, 106–148. doi:10.1111/j.1467-6494.1986.tb00391.x.

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*, 3–5. doi:10.1177/1745691610393980.

Byrne, B. M. (2012). *Structural equation modeling with Mplus: Basic concepts, applications, and programming*. New York, NY US: Routledge/Taylor & Francis Group. doi:10.1111/insr.12051_1.

Carli, M., Delle Fave, A., & Massimini, F. (1988). The quality of experience in the flow channels: Comparison of Italian and U.S. students. In M. Csikszentmihalyi & I. Csikszentmihalyi (Eds.), *Optimal*

experience: Psychological studies of flow in consciousness (pp. 288–318). NY: Cambridge University Press. doi:10.1177/0008417498065001O5.
Chalofsky, N., & Krishna, V. (2009). Meaningfulness, commitment, and engagement: The intersection of a deeper level of intrinsic motivation. *Advances in Developing Human Resources, 11*(2), 189–203. doi:10.1177/1523422309333147.
Chamberlain, K., & Zika, S. (1988). Measuring meaning in life: An examination of three scales. *Personality and Individual Differences, 9*(3), 589–596. doi:10.1016/0191-8869(88)90157-2.
Coffey, J. K., Warren, M., & Gottfried, A. (2014). Does infant happiness forecast adult life satisfaction? Examining subjective well-being in the first quarter century of life. *Journal of Happiness Studies*. doi:10.1007/s10902-014-9556-x.
Cohn, M., & Fredrickson, B. (2009). Positive emotions. In S. J. Lopez & C. R. Snyder (Eds.), *Oxford handbook of positive psychology* (2nd ed., pp. 13–24). NY: Oxford University Press.
Csikszentmihalyi, M. (1988). The flow experience and its significance for human psychology. In M. Csikszentmihalyi & I. Csikszentmihalyi (Eds.), *Optimal experience: Psychological studies of flow in consciousness* (pp. 15–35). NY: Cambridge University Press.
Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience*. New York: Harper & Row.
Csikszentmihalyi, M. (2008). *Flow: The psychology of optimal experience*. NY: Harper Perennial Modern Classics.
Debats, D., Van der Lubbe, P., & Wezeman, F. (1993). On the psychometric properties of the life regard index (LRI): A measure of meaningful life: An evaluation in three independent samples based on the dutch version. *Personality and Individual Differences, 14*, 337–345. doi:10.1016/0191-8869(93)90132-M.
Diener, E. (2000). Subjective well-being: The science of happiness and a proposal for a national index. *American Psychologist, 55*(1), 34–43. doi:10.1037/0003-066X.55.1.34.
Diener, E., & Emmons, R. A. (1985). The independence of positive and negative affect. *Journal of Personality and Social Psychology, 47*, 1105–1117. doi:10.1037/0022-3514.47.5.1105.
Diener, E., Emmons, R., Larsen, R., & Griffin, S. (1985). The satisfaction with life scale. *Journal of Personality Assessment, 49*(1), 71–75. doi:10.1207/s15327752jpa4901_13.
Diener, E., & Oishi, S. (2000). Money and happiness: Income and subjective well-being across nations. In E. Diener & E. M. Suh (Eds.), *Culture and subjective well-being* (pp. 185–218). Cambridge, MA US: The MIT Press.
Diener, E., & Seligman, M. (2002). Very happy people. *Psychological Science, 13*(1), 81–84. doi:10.1111/1467-9280.00415.
Diener, E., Wirtz, D., Tov, W., Kim-Prieto, C., Choi, D., Oishi, S., et al. (2010). New well-being measures: Short scales to assess flourishing and positive and negative feelings. *Social Indicators Research, 97*(2), 143–156. doi:10.1007/s11205-009-9493-y.
Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology, 92*(6), 1087–1101. doi:10.1037/0022-3514.92.6.1087.
Duckworth, A., & Quinn, P. D. (2009). Development and validation of the short grit scale (GRIT–S). *Journal of Personality Assessment, 91*(2), 166–174. doi:10.1080/00223890802634290.
Enders, C. K. (2010). *Applied missing data analysis*. New York: Guilford Press.
Engeser, S., Rheinberg, F., Vollmeyer, R., & Bischoff, J. (2005). Motivation, flow-experience, and performance in learning settings at Universities. *German Journal of Educational Psychology, 19*, 159–172. doi:10.1024/1010-0652.19.3.159.
Forgeard, M. J. C., Jayawickreme, E., Kern, M., & Seligman, M. E. P. (2011). Doing the right thing: Measuring wellbeing for public policy. *International Journal of Wellbeing, 1*(1), 79–106. doi:10.5502/ijw.v1i1.15.
Fredrickson, B. L. (2001). The role of positive emotions in positive psychology: The broaden-and-build theory of positive emotions. *American Psychologist, 56*(3), 218–226. doi:10.1037/0003-066X.56.3.218.
Gagné, M. (2003). The role of autonomy support and autonomy orientation in prosocial behavior engagement. *Motivation and Emotion, 27*, 199–223. doi:10.1023/A:1025007614869.
Graham, J. W. (2012). *Missing data: Analysis and design*. Berlin: Springer.
Graham, J. W., & Coffman, D. L. (2012). Structural equation modeling with missing data. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 277–295). New York: Guilford Press.
Grant, H., & Dweck, C. (2003). Clarifying achievement goals and their impact. *Journal of Personality and Social Psychology, 85*(3), 541–553. doi:10.1037/0022-3514.85.3.541.

Han, S. (1988). The relationship between life satisfaction and flow in elderly Korean immigrants. In M. Csikszentmihalyi & I. Csikszentmihalyi (Eds.), *Optimal experience: Psychological studies of flow in consciousness* (pp. 138–149). NY: Cambridge University Press.

Huppert, F. A. (2009). Psychological well-being: Evidence regarding its causes and consequences. *Applied Psychology: Health and Well-Being, 1*(2), 137–164. doi:10.1111/j.1758-0854.2009.01008.x.

Huppert, F. A., & So, T. C. (2013). Flourishing across Europe: Application of a new conceptual framework for defining well-being. *Social Indicators Research, 110*(3), 837–861. doi:10.1007/s11205-011-9966-7.

Kashdan, T. B., Julian, T., Merritt, K., & Uswatte, G. (2006). Social anxiety and posttraumatic stress in combat veterans: Relations to well-being and character strengths. *Behavior Research and Therapy, 44*, 561–583. doi:10.1016/j.brat.2005.03.010.

Kern, M., Waters, L., Adler, A., & White, M. (2014). Assessing employee wellbeing in schools using a multifaceted approach: Associations with physical health, life satisfaction, and professional thriving. *Psychology, 5*, 500–513. doi:10.4236/psych.2014.56060.

Keyes, C. (2007). Promoting and protecting mental health as flourishing: A complementary strategy for improving national mental health. *American Psychologist, 62*(2), 95–108. doi:10.1037/0003-066X.62.2.95.

King, G., & Roberts, M. E. (2014). How robust standard errors expose methodological problems they do not fix, and what to do about it. Retrieved from http://j.mp/InK5jU.

Layous, K., Nelson, S. K., & Lyubomirsky, S. (2013). What is the most optimal way to deliver a positive activity intervention? The case of writing about one's "best possible selves". *Journal of Happiness Studies, 14*, 635–654. doi:10.1007/s10902-012-9346-2.

LeFevre, J. (1988). Flow and the quality of experience during work and leisure. In M. Csikszentmihalyi & I. Csikszentmihalyi (Eds.), *Optimal experience: Psychological studies of flow in consciousness* (pp. 307–318). NY: Cambridge University Press.

Liu, S., Rovine, M. J., & Molenaar, P. C. M. (2012). Selecting a linear mixed model for longitudinal data: repeated measures analysis of variance, covariance pattern model, and growth curve approaches. *Psychological Methods, 17*, 15–30. doi:10.1037/a0026971.

Lyubomirsky, S., King, L., & Diener, E. (2005). The benefits of frequent positive affect: Does happiness lead to success? *Psychological Bulletin, 131*(6), 803–855. doi:10.1037/0033-2909.131.6.803.

Martin, L. R., Friedman, H. S., Tucker, J. S., Tomlinson-Keasey, C., Criqui, M. H., & Schwartz, J. E. (2002). A life course perspective on childhood cheerfulness and its relation to mortality risk. *Personality and Social Psychology Bulletin, 28*(9), 1155–1165. doi:10.1177/01461672022812001.

Mason, W., & Suri, S. (2012). Conducting behavioral research on Amazon's mechanical Turk. *Behavior Research Methods, 44*(1), 1–23. doi:10.3758/s13428-011-0124-6.

Muthén, L.K. and Muthén, B.O. (1998–2012). Mplus User's Guide. Seventh Edition. Los Angeles, CA: Muthén & Muthén.

Nelson, S. K., Fuller, J. A. K., Choi, I., & Lyubomirsky, S. (2014). Beyond self-protection: Self-affirmation benefits hedonic and eudaimonic well-being. *Personality and Social Psychology Bulletin, 40*, 998–1011. doi:10.1177/0146167214533389.

Nelson, S. K., & Lyubomirsky, S. (2012). Finding happiness: Tailoring positive activities for optimal well-being benefits. In M. Tugade, M. Shiota, & L. Kirby (Eds.), *Handbook of positive emotions*. NY: Guilford.

Paolacci, G., & Chandler, J. (2014). Inside the turk: Understanding mechanical turk as a participant pool. *Current Directions in Psychological Science, 23*(3), 184–188. doi:10.1177/0963721414531598.

Pennebaker, J. W. (1982). *The psychology of physical symptoms*. New York: Springer.

Peterson, C. (2006). *A primer in positive psychology*. NY: Oxford University Press.

Peterson, C., Ruch, W., Beermann, U., Park, N., & Seligman, M. P. (2007). Strengths of character, orientations to happiness, and life satisfaction. *The Journal of Positive Psychology, 2*(3), 149–156. doi:10.1080/17439760701228938.

Preacher, K. J., & Coffman, D. L. (2006). Computing power and minimum sample size for RMSEA [Computer software]. Available from http://quantpsy.org/.

Radloff, L. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement, 1*(3), 385–401. doi:10.1177/014662167700100306.

Rogatko, T. (2009). The influence of flow on positive affect in college students. *Journal of Happiness Studies, 10*(2), 133–148. doi:10.1007/s10902-007-9069-y.

Ryan, R., & Deci, E. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55*(1), 68–78. doi:10.1037/0003-066X.55.1.68.

Ryan, R., & Deci, E. (2001). On happiness and human potentials: A review of research on hedonic and eudaimonic well-being. *Annual Review of Psychology, 52*, 141–166. doi:10.1146/annurev.psych.52.1.141.

Ryan, R., & Frederick, C. (1997). On energy, personality and health: Subjective vitality as a dynamic reflection of well-being. *Journal of Personality, 65*, 529–565. doi:10.1111/j.1467-6494.1997.tb00326.x.

Ryff, C. (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. doi:10.1037/0022-3514.57.6.1069.

Sandvik, E., Diener, E., & Seidlitz, L. (1993). Subjective well-being: The convergence and stability of self-report and non-self-report measures. *Journal of Personality, 64*, 317–342. doi:10.1007/978-90-481-2354-4_6.

Seligman, M. (2011). *Flourish: A visionary new understanding of happiness and well-being*. NY: Free Press.

Seligman, M., & Csikszentmihalyi, M. (2000). Positive psychology: An introduction. *American Psychologist, 55*(1), 5–14. doi:10.1037/0003-066X.55.1.5.

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

Steger, M., Oishi, S., & Kashdan, T. (2009). Meaning in life across the life span: Levels and correlates of meaning in life from emerging adulthood to older adulthood. *The Journal of Positive Psychology, 4*(1), 43–52. doi:10.1080/17439760802303127.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 54*(6), 1063–1070. doi:10.1037/0022-3514.54.6.1063.