American Psychological Association

Psychological Assessment

ISSN: 1040-3590

2020, Vol. 32, No. 3, 294–313
http://dx.doi.org/10.1037/pas0000787

The Well-Being Profile (WB-Pro): Creating a Theoretically Based Multidimensional Measure of Well-Being to Advance Theory, Research, Policy, and Practice

Herbert W. Marsh and Felicia A. Huppert
Australian Catholic University

James N. Donald
The University of Sydney Business School

Marcus S. Horwood and Baljinder K. Sahdra
Australian Catholic University

There is no universally agreed definition of well-being as a subjective experience, but Huppert and So (2013) adopted and systematically applied the definition of well-being as positive mental health—the opposite of the common mental disorders described in standard mental health classifications (e.g., *Diagnostic and Statistical Manual of Mental Disorders*). We extended their theoretical approach to include multi-item scales, using 2 waves of nationally representative U.S. adult samples to develop, test, and validate our multidimensional measure of well-being (WB-Pro). This resulted in a good-fitting a priori (48-item, 15-factor) model that was invariant over time, education, gender, and age; showed good reliability (coefficient αs .81–.93), test–retest correlation (.73–.85; M = .80), and convergent/discriminant validity based on a multitrait-multimethod analysis, and relations with demographic variables, selected psychological measures, and other multidimensional and purportedly unidimensional well-being measures. Further, we found that items from 2 widely used, purportedly unidimensional well-being measures loaded on different WB-Pro factors consistent with a priori predictions based on the WB-Pro factor structure, thereby calling into question their claimed unidimensionality and theoretical rationale. Because some applications require a short global measure, we used a machine-learning algorithm to construct 2 global well-being short versions (five- and 15-item forms) and tested these formative measures in relation to the full-form and validity criteria (to download short and long versions see https://ippe.acu.edu.au/research/research-instruments/wb-pro). The WB-Pro appears to be one of the most comprehensive measures of subjective well-being, based on a sound conceptual model and empirical support, with broad applicability for research and practice, as well as providing a framework for evaluating the breadth of other well-being measures.

Public Significance Statement
Based on a systematic and coherent theoretical approach to defining well-being, we have used state-of-the-art psychometric techniques to develop and validate a new multidimensional well-being measure (WB-Pro). The full professional version (48-items, 15-factors) is recommended for high-quality well-being research. It is also available in shorter versions (15 and 5-items) that can be utilized where is it not practical to use the full version (to download short and long versions see https://ippe.acu.edu.au/research/research-instruments/wb-pro).

Keywords: multiple dimensions of well-being, convergent and discriminant validity, formative, reflective, and unidimensional scales, factor analysis, machine-learning

Supplemental materials: http://dx.doi.org/10.1037/pas0000787.supp

There are many approaches to conceptualizing well-being, and each has different implications for how well-being should be measured. Some researchers use the term well-being very broadly to include objective circumstances, including economic and social

This article was published Online First December 12, 2019.
Herbert W. Marsh and Felicia A. Huppert, Institute for Positive Psychology and Education, Australian Catholic University; James N. Donald, The University of Sydney Business School; Marcus S. Horwood and Baljinder K. Sahdra, Institute for Positive Psychology and Education, Australian Catholic University.

Correspondence concerning this article should be addressed to Herbert W. Marsh, Institute for Positive Psychology and Education, Australian Catholic University, Level 10, 33 Berry Street, North Sydney NSW 2060, P.O. Box 968, North Sydney, NSW 2059. E-mail: herb.marsh@acu.edu.au

conditions (e.g., income, housing, and education) or lack of mental ill-health symptoms (e.g., depression and anxiety), but our focus is on well-being as a subjective experience beyond simply neutral levels of mental ill-health. We briefly review the many different approaches to conceptualizing well-being in its subjective sense to set the context for the present study.[1]

Conceptual Framework for Defining Well-Being

At its broadest level, well-being refers to our perception of how well our life is going. For some researchers, this concept is best captured by a global evaluation of how satisfied we are with our life. This evaluation most frequently relies on a single question, and it appears the earliest reference to such a survey measure is Cantril (1965). Some scholars prefer to use several questions to evaluate life satisfaction (e.g., Cummins, 1996; Diener, Emmons, Larsen, & Griffin, 1985). For certain purposes, this global evaluation, particularly when it results in a single number, can be very useful. For example, economists are increasingly interested in measuring subjective well-being in national and cross-national surveys to supplement the traditional reliance on economic indicators as the principal drivers of policy. In this context, they argue that a single numerical estimate of subjective well-being is necessary to place alongside a single numerical estimate of measures such as GDP (e.g., Clark, Flèche, Layard, Powdthavee & Ward., 2018).

In contrast, some scholars view well-being as the presence of positive emotions such as happiness, rather than as a global evaluation of life (e.g., Bradburn, 1969; Fredrickson, 2009; Kahneman, Diener, & Schwarz, 1999; Layard, 2005). This is the hedonic view of well-being. Among the earliest scales that measure this construct are Bradburn's Affect Balance Scale (Bradburn, 1969) and the Positive and Negative Affect Scale (PANAS) of Watson, Clark, and Tellegen (1988), while the Scale of Positive and Negative Experience (SPANE) of Diener et al. (2010) is among the most recent. One of the critiques of the view that well-being can be equated with positive affect is that emotions are ephemeral, whereas a longer perspective is needed to establish how well our life is going. In addition, happiness can be achieved in ways that are very unhelpful to the individual in the longer term, such as the use of mood enhancing drugs.

Another conceptual approach is that well-being cannot be reduced to either happiness or life satisfaction alone, but that well-being comprises a number of different components (Huppert & Ruggeri, 2018). According to Diener, Suh, Lucas, and Smith (1999), "subjective well-being" is the combination of life satisfaction, the presence of pleasant affect and the absence of unpleasant affect. Other scholars postulate that the subjective experience of well-being is more than the combination of feeling good and being satisfied; it also includes functioning well both personally and socially. This is sometimes referred to as eudaimonic well-being (Ryan & Deci, 2001). Some scholars conceptualize well-being purely in terms of eudaimonic well-being. This includes Ryff's (1989) Psychological Well-being, which describes six dimensions of personal and social functioning: autonomy, environmental mastery, personal growth, positive relationships, purpose in life, and self-acceptance. Another example of a eudaimonic conceptualization is self-determination theory (Ryan & Deci, 2000), which postulates that the fulfillment of three basic psychological needs—autonomy, competence, and relatedness—is both essential and sufficient for well-being. However, a number of leading authorities (e.g., Keyes & Waterman, 2003; Seligman, 2002, 2011) argue that true well-being requires the combination of hedonic well-being (positive affect) and eudaimonic well-being. This framework has also been adopted by the Organization for Economic Cooperation and Development (OECD), which regularly undertakes cross-national well-being surveys. Their guidelines state that well-being measures should include positive affect, eudaimonic well-being, and life evaluation (OECD, 2013).

In a pragmatic approach to creating a consensus on the components of well-being, Su, Tay, and Diener (2014) undertook a review that combined all the existing approaches. They describe a multicomponent model with 18 facets of positive functioning, representing seven dimensions of psychological well-being: subjective well-being (satisfaction with life, positive emotions), relationships (positive relations with others, belonging), meaning (purpose in life), engagement, mastery (competence, self-efficacy, and self-esteem), optimism, and autonomy (need for autonomy, control). Despite this pragmatic effort of Su et al. (2014), there remains a lack of a consensus on an appropriate theoretical framework.

One theoretical framework with a long history defines psychological well-being as equivalent to positive mental health, and includes both eudemonic and hedonic aspects of well-being. This approach has been adopted by many organizations and individuals, including the World Health Organization (2006), where mental well-being was equated with mental health and the term, "mental health" was used to refer to a positive state, and not just the absence of mental disorders. Jahoda (1958) contrasted psychological well-being with psychological ill-being and defined six elements of positive psychological functioning: attitudes of an individual toward his or her own self, personal growth or self-actualization, integration, autonomy, perception of reality, and environmental mastery. A similar approach was taken by Ryff (1989), who also proposed six dimensions of psychological well-being.

The present study builds on the theoretical approach that defines well-being as equivalent to positive mental health. One of the benefits of the positive mental health approach is that, in contrast to the lack of agreement on the components of psychological well-being, there is international agreement on the components (symptoms) of psychological ill-being as listed in the International Classification of Diseases (ICD) Mental and Behavioral Disorders

[1] The literature contains some unhelpful terminology, with some authors using the terms "subjective well-being" and "psychological well-being" in very restrictive ways and strictly differentiated from each other. However, because our focus is on well-being in its general subjective sense, and because each of these terms actually refers to experiences that are both subjective and psychological, we use the terms interchangeably. On the occasions when we use the term "well-being" on its own, it refers to these subjective aspects of well-being as opposed to objective aspects of well-being. It is worth noting that there is some overlap between the terms well-being and "quality of life," although the latter is typically used in the context of healthcare. Quality of life scales may include questions about how people feel and aspects of their psychological functioning, but in contrast to well-being measures, they emphasize health status, physical capability, and ability to function in a sociocultural context (WHOQoL Group, 1994).

(World Health Organization, 1990, 2018) and the *Diagnostic and Statistical Manual of Mental Disorders* (*DSM*; American Psychiatric Association, 1994, 2013). The basic lists of symptoms of mental disorders remain virtually unchanged across different editions of these manuals, even though there is continued debate on whether and how individual diseases should be classified (Kotov et al., 2017). We acknowledge that we did not include symptoms from all mental disorders that could have broadened the model. However, our focus was on the disorders which are regarded as common mental disorders not only in the sense of their high prevalence, but also in the sense that they can affect anyone in the general population, which is not the case with most other mental disorders.

Building on this body of knowledge about the symptoms of ill-being, Huppert and So (2013) suggested that well-being goes beyond a neutral point that merely reflects an absence of mental ill-being symptoms, proposing a typology of well-being that is the opposite of the symptoms of the common mental disorders. They began with a list of the psychological symptoms and criteria used to describe the most common mental disorders, depression and anxiety, as described in the two widely used international classifications of diseases (*DSM*, ICD). They focused on categories of Major Depressive Episode (*DSM*) and Depressive Episode (ICD), and on Generalized Anxiety Disorder (terminology common to both systems) that were most prevalent and had logical polar opposites. Conceiving well-being as lying at the opposite end of the spectrum to the common mental disorders, Huppert and So (2013) identified features such as happiness and hopefulness (the opposite of specific depression symptoms) or calmness and resilience (the opposite of core symptoms of generalized anxiety). In addition, they included *DSM* Axis V-Global Assessment of Functioning, which rates an individual's general level of personal and social functioning, and which the ICD classification includes within their diagnostic criteria.

This systematic approach yielded 10 features of positive well-being. It was found that these 10 features combined positive feeling and positive functioning (i.e., hedonic and eudaimonic aspects of well-being): competence, emotional stability, engagement, meaning, optimism, positive emotion, positive relationships, resilience, self-esteem, and vitality.

Framework for Measuring Well-Being

As can be seen from the preceding brief review of the different approaches to conceptualizing well-being, there seems to be widespread agreement that well-being is best understood as a multidimensional construct, so measures of well-being need to reflect this. Of course, it would be convenient if we could measure how well a person perceives their life to be going using a single construct or better still, a single question such as a life satisfaction or happiness question, but this would not be very illuminating if well-being is really a multidimensional construct. A single construct or single item could not provide useful information about the profile of different components that make up well-being, and as a result, it could not provide practical guidance to policymakers about which components of well-being need to be improved, nor specific interventions to improve these components. The need for a multidimensional approach to measuring well-being is evident in large scale studies showing that groups or countries can obtain identical scores on a life satisfaction measure, but display completely different profiles on well-being dimensions (Huppert & So, 2013).

A number of scales specifically recognize the multidimensional nature of well-being and provide subscales for measuring different dimensions. This includes the PERMA Profiler (Butler & Kern, 2016) based on Seligman's (2011) five proposed pillars of well-being (positive emotion, engagement, relationships, meaning, and accomplishment), Ryff and Singer's (1996) Psychological Well-being scales, and Ryan and Deci's (2000) Basic Psychological Needs Satisfaction Scale, as well as the Comprehensive Inventory of Thriving (CIT) of Su et al. (2014).

Other popular scales, although not providing subscales for measuring specific dimensions, have nevertheless recognized the multidimensional nature of well-being, by including multiple items that cover a range of well-being constructs. These include, the Warwick-Edinburgh Mental Well-being Scale of Tennant et al. (2007) and the Flourishing Scale of Diener et al. (2010). However, the way in which these scales are scored usually results in a single, total score. While it is important to use a range of items to capture the diversity of processes and characteristics that underpin well-being, converting these items into a single score does not provide insight into how an individual is doing on the various dimensions of well-being. As emphasized by Marsh (2007), if a survey instrument is an ill-defined mix of different items that are not supported by a well-established factor structure and are summarized by an average of these items, then there is no basis for knowing what is being measured. Here we take an alternative approach, starting with a systematic conceptual framework for measuring well-being.

Huppert and So (2013) developed an operational definition of flourishing based on psychometric analysis of indicators of 10 components of well-being, using data from well-being questions that were administered to a representative sample of 43,000 Europeans who participated in Round 3 of the European Social Survey (ESS, Huppert et al., 2009). There were striking differences in country profiles across the 10 features. Huppert and So concluded that their profiles offered fresh insights into cultural differences in well-being as well as promising targets for policies to improve well-being. Their comparison with a life satisfaction measure showed that valuable information would be lost if well-being had been measured only by life satisfaction. Taken together, their findings reinforce the need to measure subjective well-being as a multidimensional construct in future surveys. However, as previously mentioned, one weakness in their study is that each component/dimension is represented by only a single item.

The Present Investigation

Components of Well-Being

The overarching purpose of our research is to develop a concise measure of well-being that produces reliable and valid scores on each of the theoretically derived dimensions of well-being, with particular emphasis on those proposed by Huppert and So (2013). Following Gogol et al. (2014) and others (e.g., Marsh, Ellis, Parada, Richards, & Heubeck, 2005), we argue that in relation typically limited testing time and cost effectiveness for large surveys, brief scales of three or four items are desirable as long as psychometric support is strong. Based on our review of the well-being literature, we identified five important constructs that had

not been included in the original list. Three of these, competence, self-acceptance, and autonomy, have an individual focus like the 10 original constructs, while two of the new components (empathy and prosocial behavior) have an interpersonal focus. The importance of these additional constructs for positive mental health was independently supported through consultations with several clinical psychologist colleagues.

The term "competence" was included in Huppert and So's (2013) original list of constructs, but it could more accurately have been described as "clear thinking," because the *DSM*/ICD criteria from which it was derived concern the ability to think, concentrate, and make decisions. In the current version, we added a more traditional measure of competence arguably a core component of general well-being and thriving (Marsh, Martin, Yeung, & Craven, 2017; Ryan & Deci, 2017). People who feel a sense of general competence have higher self-esteem (Thøgersen-Ntoumani & Ntoumanis, 2007), and greater satisfaction with life (Meyer, Enström, Harstveit, Bowles, & Beevers, 2007). Conversely, individuals with anxiety and depression have difficulty achieving goals and report feeling a lack of general competence (e.g., Ryan & Deci, 2017; Wei, Shaffer, Young, & Zakalik, 2005).

Considerable theory and research based on the SDT (Ryan & Deci, 2001, 2017) argues a lack of autonomy underpins all of the common mental disorders. Depression and anxiety are associated with decrements in perceived volition and control over one's life, and the tendency to make decisions out of shame, guilt, or avoidance, rather than one's longer term values and aspirations (Ryan & Deci, 2017). All of these tendencies are linked to an absence of autonomy, meaning a general sense of autonomy is a core component of healthy functioning (Ryan & Deci, 2017).

Further, depression and anxiety are linked to a general nonacceptance of oneself and one's life (Hayes, Strosahl, & Wilson, 1999; Hayes, Strosahl, & Wilson, 2012). In clinical and nonclinical samples, nonacceptance is strongly correlated with measures of general psychopathology (Hayes et al., 2004) and specific measures of anxiety and depression (Forsyth, Parker, & Finlay, 2003; Marx & Sloan, 2005; Roemer, Salters, Raffa, & Orsillo, 2005; Tull, Gratz, Salters, & Roemer, 2004). Indeed, third-wave behavioral theories of psychopathology such as Acceptance and Commitment Therapy have identified the tendency toward the avoidance and nonacceptance of one's internal states (i.e., thoughts and emotions) as a generalized process underlying all psychopathology (Hayes et al., 1999). In this way, self-acceptance is a core feature of subjectively experienced well-being, wherein one's internal states are acknowledged and not suppressed.

Empathy and prosocial behavior were included on the basis that prosocial emotions and behaviors are central to human functioning and vitality (Eisenberg, Fabes, & Spinrad, 2007; Weinstein & Ryan, 2010). Empathy is the tendency to vicariously experience other individuals' emotional states (Davis, 1994). Individuals with some mental health disorders have difficulty feeling the emotions of others and taking the perspective of others (e.g., Baron-Cohen, 2011; Caligor, Levy, & Yeomans, 2015). In contrast, empathy is essential to positive social functioning (Batson, 1991; Eisenberg et al., 2007) and has been associated with group cohesion (Henry, Sager, & Plunkett, 1996), relationship satisfaction (Davis & Oathout, 1987), and prosocial behavior (Sahdra, Ciarrochi, Parker, Marshall, & Heaven, 2015), and as such is an important feature of healthy individual functioning.

Prosocial behavior has been defined as "voluntary behavior intended to benefit another" (Eisenberg et al., 2007, p. 646). It is related to empathy but conceptually distinct from it, in that the former describes observable behavior, whereas the latter describes an internal state. Depression and anxiety have also been negatively linked to prosocial behavior, including social withdrawal, and less capacity to respond to the needs of others (Eisenberg et al., 2007).

Based on the above considerations, it was decided to add these five constructs to the 10 described by Huppert and So (2013). Although not directly derived from the inverse of *DSM*/ICD classifications, all are at least indirectly related to these classifications, and vetted by clinical psychologists. This provides an opportunity to examine how the original constructs and the newer constructs are related to each other, how each is related to a standard measure of well-being such as life satisfaction, and to test the independence of the 15 constructs.

WB-Pro: Development and Refinement of a Preliminary Item Pool

The preliminary development and refinement of an item pool took place in three stages. In the first stage, an extensive pool of a total of 195 items was prepared by the authors to represent the 15 WB-Pro factors—the 10 originally proposed by Huppert and So (2013) and the additional five factors added in the present investigation. In an initial evaluation, 29 leading academic colleagues with relevant backgrounds were provided with the 15 constructs and sorted each of 195 items into what they considered to be the most appropriate construct (or constructs) and commented on the suitability of each item. On the basis of this feedback, we refined the item pool to a total of 132 items representing the 15 constructs. A Qualtrics survey was then developed in which the order of items was randomized for each participant and administered electronically to a large, representative sample of U.S. adults. Using the standard procedures for developing a "short form" (see Marsh et al., 2005; Marsh, Martin, & Jackson, 2010), the item pool used at Time 1 (T1) was reduced to an item pool of 60 items to represent the 15 constructs, and these were the basis of the survey administered at Time 2 (T2). In particular, items were retained that: loaded substantially on the factor they were designed to measure; contributed to the reliability of the scale; did not cross-load substantially on other scales; and did not have substantial correlated uniquenesses with other items. Care was also taken to select items that maintained the breadth of the original construct. Using a similar procedure, the set of 60 WB-Pro items retained at T2 was further reduced to the 48 items comprising the final version of the WB-Pro. All analyses presented here are based on this final set of 48 items (see Figure 1).

In summary, based on these selection procedures and traditional criteria of a psychometrically sound instrument, we sought to construct the WB-Pro instrument such that it demonstrates:

- Good reliability: Median Cronbach's coefficient $\alpha \geq .80$ across the scales (T1 and T2);
- Good test–retest stability over 3 months: median test–retest correlation $\geq .70$ across the 15 scales (repeat sample from T1 and T2);
- A well-defined, replicable factor structure as shown by structural equation modeling in relation to traditional indices of fit (Marsh, Hau, & Wen, 2004; T1 and T2);

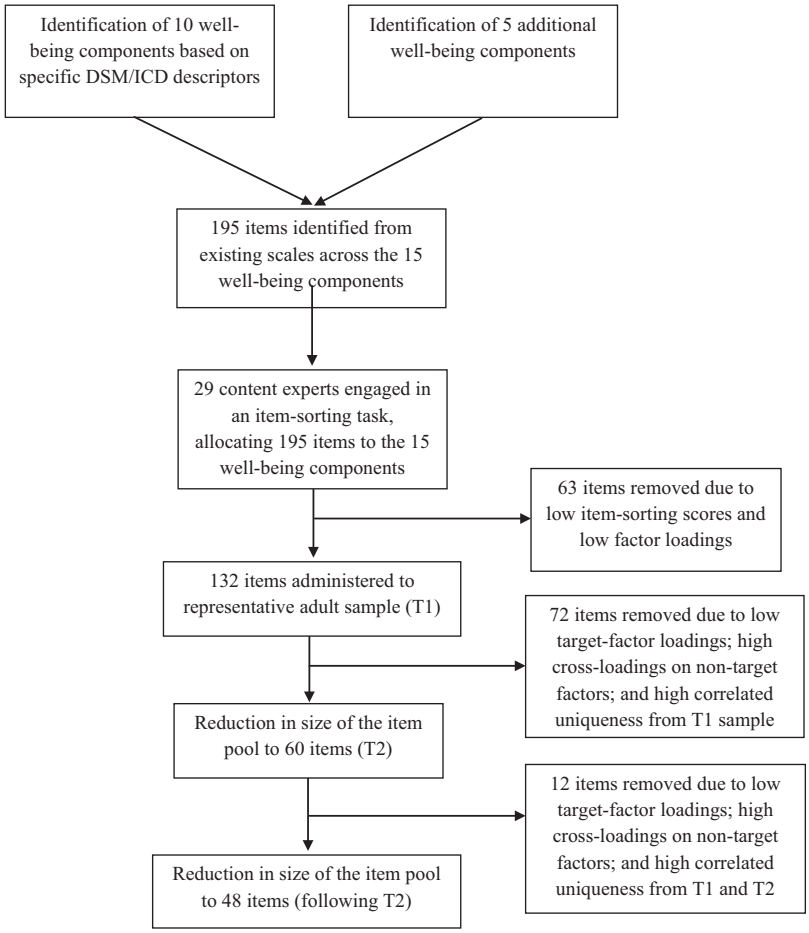

Figure 1. Flow diagram of processes for selection of items used in the well-being profile (WB-Pro) 48.

- A factor structure that is invariant over gender, age, level of education, and time as shown by multiple-group structural equation models (T1 and T2);
- Applicability for participants across the age range from late-adolescent/young adult, middle-age, and older adults (combined sample from T1 and T2);
- Convergent and discriminant validity in relation to: (a) multitrait-multimethod (MTMM) analyses of WB-Pro responses in relation to time (test–retest stability, T1 and T2); (b) other multidimensional measures of well-being (e.g., PERMA, Basic Psychological Needs, T2); (c) multi-item, purportedly unidimensional well-being instruments (e.g., WEMWBS, Flourishing, T2); (d) other selected psychological measures (depression, stress; life satisfaction, happiness, sleep, and life-event changes); and selected demographic variables (gender, age, marital status, English fluency, and education).

Finally, we recognize that some applications require a short global measure rather than the full multidimensional WB-Pro. This might be because particular applications might not be able to include the full 48-item measure because of time constraints, or because researchers are only interested in a global measure of well-being. For this reason, we used a novel machine learning approach to construct two global well-being short-forms. The 15-item version (WB-Pro15), maintains the full range of dimensions by selecting the best item from within each of the 15 dimensions. The even shorter five-item version (WB-Pro5) selects the best five items from among the 48-items. We offer these short, global measures as formative indices of well-being, but emphasize that the full version WB-Pro provides a more reliable and robust, multidimensional representation of well-being.

Further, we note there has been considerable theoretical and statistical confusion in the construction and evaluation of global measures of well-being based on single-item indicators of selected components of well-being, as in the 14-item WEMWBS (Tennant et al., 2007), and eight-item Flourishing Scale (Diener et al., 2010). It is our contention that these multi-item global measures, like the WB-Pro5 and WB-Pro15, should be considered formative rather than reflective measures, and that psychometric criteria like unidimensionality are inappropriate and, perhaps, even counterproductive in the construct and evaluation of formative scales. In support of these claims, we demonstrate that WEMWBS and Flourishing measures really are multidimensional measures. First, we conceptually map the 14 WEMWBS items and the eight Flourishing items onto the 15 WB-Pro factors based on item content. Then we provide empirical support for this a priori con-

ceptual mapping based on factor analyses of all 60 items (14 WEMWBS, 8 Flourishing, and 48 WB-Pro). Implications of this distinction between formative and reflective measures are then discussed.

Method

Participants

Data were collected from a nationally representative sample of U.S. adults sourced by Qualtrics, an enterprise survey technology solution entity. Ethical approval for the study was granted by the Research Ethics Board (2018-288H and 2018-289H). Qualtrics sourced the participants from several traditional and online market research panels and provided with a small monetary benefit for completing the surveys (between $2–5 USD worth of points that could be traded for merchandise). Participants were provided with a unique survey Web link to allow for them to be identified and matched for completed surveys in T1 and T2. Qualtrics used preliminary data screening to ensure that there were not multiple responses by the same respondent on the basis IP addresses and other preliminary data screening in relation to completion of the survey. Based on the data provided by Qualtrics, we used preliminary data screening filters to eliminate questionable responses: two validity check items instructing respondents to select a specific response; and a visual check for response patterns such as straight-lining or (i.e., selecting the same response in a row) or diagonal-lining. On this basis 3.2% of the respondents were eliminated.

As part of the first survey (T1), participants were also asked to indicate whether they would wish to be involved in the follow up survey (T2). The T2 survey was conducted 3 months after the T1 survey. Participants completed an online anonymous survey in exchange for points they received from the survey company, which could then be redeemed for merchandise. All participants provided basic demographic information (i.e., age, gender, marital status, education level, employment status, English fluency, and ethnic background). To increase the likelihood of the collected sample being representative of the general U.S. population, quota aims were established to align with the U.S. population breakdown of age and gender demographics, based on the figures published by the U.S. Census Bureau, with the samples collected for this study. There was minimal difference between the demographic breakdowns of the samples we used in both T1 and T2 and the national demographic data published by the U.S. Census Bureau.

A total of 1,035 participants completed the online survey at time T1; 51% female with a mean age of 45.5 years (SD = 16.5). Of these 1,035 participants, 444 also completed the T2 online survey, in addition to 1,082 newly sourced participants, making a total of 1,524 participants for T2; 49% female participants with a mean age of 47.7 years (SD = 16.9). In total, we considered 2559 sets of responses by 2,117 participants that we divided into three groups: 1 (N = 593 surveys by those completing surveys at T1 only); 2 (N = 1082 surveys by participants completing surveys at T2 only), and 3 (N = 884, surveys by participants completing surveys at T1 and T2). Different sets of analyses were based on different groups of respondents.

Measures: Key Correlates and Determinants of Our 15 Well-Being Factors

Life changes. As part of the surveys at both T1 and T2, participants were asked: "Please indicate whether you have experienced any major changes in your life, whether positive or negative, during the past three months." Across the 2,559 sets of responses, only 24% indicated that there had been a major change. We then classified these as negative (10.3%), neutral (5.3%), positive (6.1%), or no significant events (78%; i.e., respondents who did not list a significant life change). For present purposes we constructed four dummy variables (with no significant events being the "left out" category) and regressed these on the 15 well-being factors such that coefficients represent the effect a negative, neutral, or positive life change (compared with not having a significant life change).

Alternative measures and correlates of well-being. At T2, we added a set of psychological measures designed to test the convergent and discriminant validity of interpretations of the WB-Pro factors: well-being, measured using the standard life satisfaction question from the U.K. Office for National Statistics (Self, Thomas, & Randall, 2012), the Warwick Edinburgh Mental Well-Being Scale (WEMWBS; Tennant et al., 2007), the Flourishing Scale (Diener et al., 2010), the PERMA-Profiler (Butler & Kern, 2016); basic psychological need satisfaction and thwarting, measured with the Psychological Need Satisfaction and Frustration Scale (Chen et al., 2015); personality, measured via the Big Five Personality Inventory (see Marsh, Lüdtke, et al., 2010; Marsh, Nagengast, & Morin, 2013); psychological stress, measured with the "stress" items of the Copenhagen Psychosocial Questionnaire II (Dicke et al., 2018; Pejtersen, Kristensen, Borg, & Bjorner, 2010); depression, measured via the Center for Epidemiologic Studies Depression scale (eight-item CES-D; Radloff, 1977; Steffick, 2000); sleep quality measured using items from the Sleep Quality Scale (Cappelleri et al., 2009); and general health adapted from a question from the European Social Survey (ESS, Huppert, Marks, et al., 2013).

Statistical Analysis

Factor analysis. In addition to traditional descriptive statistics and reliability estimates, factor analysis was our primary statistical tool. As noted in their classic Annual Review of Clinical Psychology article, Marsh, Morin, Parker, and Kaur (2014) emphasized that exploratory factor analysis (EFA) and confirmatory factor analysis (CFA), path analysis, and structural equation modeling (SEM) have long histories in psychological research. Although CFA has seemed to largely supersede EFA, CFAs of multidimensional constructs typically fail to meet standards of good measurement: goodness of fit, measurement invariance, lack of differential item functioning, and well-differentiated factors in support of discriminant validity. Part of the problem is undue reliance on overly restrictive CFAs in which each item loads on only one factor. Using both "real" and simulated data, previous research has shown that when this assumption of CFA is violated, CFA factor correlations tend to be positive biased and that in some cases this bias can be substantial (Marsh et al., 2014; also see Marsh et al., 2019; Morin, Marsh, & Nagengast, 2013). Exploratory SEM (ESEM), an overarching integration of the best aspects of CFA/

SEM and traditional EFA, provides confirmatory tests of a priori factor structures, and relations between latent factors and multigroup/multioccasion tests of full measurement invariance (e.g., configural, metric, and scalar invariance). Due in part to the bias in CFA when assumptions that each item loads on only a single factor, ESEM tends to result in better differentiation among the multiple factors. It incorporates all combinations of CFA factors, ESEM factors, covariates, grouping/multiple-indicator multiple-cause (MIMIC) variables, latent growth, and complex structures that typically have required CFA/SEM. Thus, ESEM has broad applicability to psychological research that is not appropriately addressed either by traditional EFA or CFA/SEM. For present purposes, the 2,559 sets of responses are based on participants who completed the survey only at Time 1 (N = 593), only at Time 2 (N = 1,082), or both Times 1 and 2 (N = 884 sets of responses by 442 individuals). To maximize the number of cases, some of the factor analyses was done for the entire set of 2,559 responses in which that data was in the "long" (or stacked) format. For these long format analyses, because 442 participants contributed responses at both T1 and T2, we used the Mplus complex design option to adjust standard errors for the fact there were two sets of responses for these participants (Muthén & Muthén, 2017). However, to evaluate the test–retest stability of responses and the invariance of the factor structure over time, we structured the data into wide format.

In the present investigation, we combined and compared CFA and ESEM solutions in the selection of items to be used in the final WB-Pro, testing psychometric properties, test–retest stability, and MTMM analyses of convergent and discriminant validity. A detailed presentation of the application of ESEM and its extension is beyond the scope of this study (but see Marsh et al., 2014 for an overview of the wide applicability of ESEM; also see Marsh, Lüdtke, et al., 2010; Marsh et al., 2009; Marsh, Lüdtke, Nagengast, Morin, & Von Davier, 2013; also see online supplemental materials, Section 1). Particularly, the application of ESEM for the development of a short form is apparently a new application of ESEM that will provide an important methodological contribution to instrument construction more generally. Assessment of goodness-of-fit was based on indices that are relatively independent of the sample-size (Hu & Bentler, 1999; Marsh, Hau, & Grayson, 2005; Marsh et al., 2004). The population values of Tucker-Lewis Index (TLI) and comparative fit index (CFI) vary along a 0-to-1 continuum, such that values greater than .90 and .95 typically reflect acceptable and excellent fits to the data, respectively. Values smaller than .08 and .06 for the root mean square error of approximation (RMSEA) support acceptable and good model fits, respectively.

Convergent and discriminant validity: MTMM analyses. The MTMM (see Campbell & Fiske, 1959) design is used widely to assess convergent and discriminant validity, and also is a standard criterion for evaluating psychological instruments. The MTMM design provides a particularly strong approach to evaluating stability of responses to a multidimensional instrument, as emphasized by Campbell and O'Connell (1967) who specifically operationalized the multiple methods in their MTMM paradigm as multiple occasions. Marsh (Marsh et al., 2005, 2010) also recommended this approach to evaluate support for the convergent and discriminant validity in relation to temporal stability over time. In this regard, convergent validities refer to stability over time (i.e., test–retest correlations) and the "method" factor is time. Although the design might be considered weak in relation to providing support for convergence based on maximally different methods (e.g., multiple respondents—self, peer; multiple instruments designed to measure the same traits), it provides a "best case" test in relation to discriminant validity. Thus, if there is no support for discriminant validity in relation to convergent validities based on time as the method factor, support for discriminant validity is unlikely to be found with other, more demanding tests of convergent validity.

Based on test–retest data from Time 1 and 2 (N = 442 participants), we assessed convergent and discriminant validity in relation to time from a within-network perspective. For these analyses, "convergent validity" is test–retest correlation, whereas the different methods refer to time (see Marsh et al., 2010 for an example of this approach). Marsh et al. (2010) demonstrated new and evolving latent-variable approaches that allow convergent and discriminant validity to be assessed using the traditional Campbell and Fiske (1959) criteria (the original and most widely used basis of assessing MTMM data) while still overcoming subsequent criticisms of these criteria. Further rationale for the MTMM approach we used in the present analyses appears in Section 4 of online supplemental materials.

Convergent and discriminant validity: Relations with other constructs. In subsequent analyses, we added demographic variables (marital status, gender, age, education, and English fluency) and covariates to the WB-Pro factor structure. The covariates included multidimensional (PERMA, Psychological Need Satisfaction, and Frustration Scale) and purportedly unidimensional measures of well-being (Flourishing, WEMWBS), as well as other measures (Big-Five personality, stress, depression, life satisfaction, happiness, sleep problems, general health, and exercise). The multidimensional measures of well-being included specific components of well-being that were included on the WB-Pro, providing a basis for testing convergent (agreement on matching factors) and discriminant (correlations for nonmatching factors) validity.

Of particular interest were the multi-item purportedly unidimensional measures of well-being—in this study, we focused on the Flourishing Scale (Diener et al., 2010) and WEMWBS (Tennant et al., 2007) measures. We began by relating these measures to our WB-Pro factors and testing the claims that these measures are unidimensional. However, the nature of the items and the basis of their selection in the original construction of these measures suggested that items were included to reflect different components of well-being. Each of the coauthors of the present study undertook a priori classification of the 22 items (eight from the Flourishing, 14 from the WEMWBS) as being associated with one or more of the 15 WB-Pro classifications. These a priori classifications were then used to test ESEM models in which the 22 items from the Flourishing and the WEMWBS were "absorbed" into the WB-Pro ESEM using target rotation in which the target factor loadings were specified based on a priori classifications (i.e., items hypothesized to load on a particular factor were given a target of .8 and all others were given targets of zero). This 15-factor solution, with the Flourishing and WEMWBS items absorbed into the WB-Pro15-factor structure, was compared with separate factor structures in which Flourishing and/or WEMWBS items were used to define separate factors (i.e., with items not cross-loading on the WB-Pro factors). The goodness-of-fit and factor loadings based on

these alternative models provided a more rigorous test of the unidimensionality of the Flourishing and the WEMWBS but also provided further tests of the convergent and discriminant validity of the WB-Pro factors.

Development of a short-form of WB-Pro. Recognizing the usefulness of a much briefer measure of well-being than the WB-Pro (based on 48 items), we undertook some additional analyses. Here the intent was to develop a formative measure of well-being based on a single global measure rather than a multidimensional profile of distinct factors. We applied a novel machine-learning approach based on genetic algorithms (GA) to the selection of the best items to reflect the measure. More specifically, we implemented the GA method in R, an open source statistical computing environment (R Core Team, 2018), using the GAabbreviate package (Scrucca & Sahdra, 2015; for further discussion see online supplemental materials, Section 8). The GA implement the principles of biological evolution (e.g., mutation, crossover, and selection based on fitness) in a computational framework to find a suitable short form of the long form that is reliable, valid, and preserves most of the variance in the data of the original questionnaire (Sahdra, Ciarrochi, Parker, & Scrucca, 2016; Yarkoni, 2010). The GA have been used to abbreviate long forms of several psychological constructs, including personality traits (Yarkoni, 2010), psychopathy (Eisenbarth, Lilienfeld, & Yarkoni, 2015), experiential avoidance (Sahdra et al., 2016), body image (Basarkod, Sahdra, & Ciarrochi, 2018), and mindfulness in sports (Noetel, Ciarrochi, Sahdra, & Lonsdale, 2019). For present purposes, we constructed two versions of WB-Pro-Short, one based on the best five items, and another based on the best 15 items subject to the constraint that at least one item was included from each of the 15 WB-Pro factors (To download short and long versions see https://ippe.acu.edu.au/research/research-instruments/wb-pro for long and short versions of the instrument).

Results

Factor Structure

CFA and ESEM. A critical initial step was to evaluate the factor structure underlying the responses to 48-item WB-Pro instrument, and to compare results based on CFA and ESEM.

Factor structure: Total group. Two sets of factor analyses–CFA and ESEM–were conducted on the entire set of 2,559 responses from participants at T1 and T2. Critical features of these analyses were the goodness-of-fit indices (see online supplemental materials, Section 3, Table S2, Models 1A and 1B). The highly restrictive CFA structure in which each item was allowed to load on one and only one factor provided a remarkably good fit (CFI = .97, TLI = .96, RMSEA = .037) in relation to traditional criteria for a good fitting model. Nonetheless, the fit of the less restrictive ESEM was even better (CFI = .99, TLI = .99, RMSEA = .023). However, the ESEM (711 parameter estimates) was less parsimonious that the CFA (249 parameter estimates). Nevertheless, even goodness-of-fit indices that control for parsimony (RMSEA and TLI) were better for the ESEM than the CFA.

Based on goodness-of-fit, both the ESEM and CFA solutions were good, with the ESEM solution having best fit, and the CFA solution being most parsimonious. Further, parameter estimates based on the CFA and ESEM both demonstrate that the WB-Pro factors are well defined. Together, both CFA and ESEM solutions provide support for the a priori factor structure relating the 48 items to the WB-Pro factors. A complete presentation of these fit statistics and parameter estimates, is in online supplemental materials, Section 3, Tables S2, S3A, and S3B.

Multiple group tests of factorial invariance. In the initial tests of invariance (invariance over stacked groups; Models 2A–2C in online supplemental materials, Section 3, Table S2), we evaluate whether the factor structures for these three groups completing WB-Pro at T1, at T2, or at both T1 and T2, are invariant. Because these groups are each nationally representative samples from the same population, it is not surprising that there is good support for even the most restrictive model of scalar invariance (i.e., invariance of factor loadings and intercepts) as well as the less restrictive models of metric (i.e., invariance of factor loadings) and configural invariance (i.e., no invariance constraints). Because scalar invariance is substantially more parsimonious than the metric and particularly the configural model, support for scalar invariance is particularly strong based on the TLI and RMSEA indices that control for parsimony.

In the next three sets of invariance tests (Models 3–5 in Table S2, online supplemental materials), we evaluate the invariance of the factor solution over three educational categories ("high school," "some tertiary education," and "a 4-year tertiary degree"), four age groups, and two gender groups (males and females). Each of these grouping variables are substantively different, and tests of invariance provide potentially demanding tests of the generalizability of the WB-Pro factor structure in relation to these demographic variables. Nevertheless, the patterns of results for each of these tests of invariance are quite similar. In each case, there is good support for even the most restrictive model of scalar invariance (Models 3C, 4C, and 5C in Table S2, online supplemental materials), as well as the less restrictive models of metric and configural invariance. Again, this support for invariance is particularly strong for the TLI and RMSEA indices that control for parsimony. Although presented only briefly here, it is important to emphasize that results of these tests of invariance demonstrated that the WB-Pro factor structure is very robust.

Convergent and Discriminant Validity in Relation to Time: MTMM Analyses

As noted earlier, that the MTMM design in relation to time provides a best case test of the discriminant validity of a multidimensional measure. Furthermore, many of the traditional problems with the original criteria proposed by Campbell and Fiske (1959) are overcome when they are applied to a latent correlation matrix based on a well-fitting factor analysis in which each factor is based on multiple items. The MTMM analysis starts with an ESEM of the 15 WB-Pro factors administered at T1 and T2. The fit of the model was good, showing strong support for invariance (configural, metric, and scalar) over time (see Models 6A, 6B, and 6C in Table S2 of online supplemental materials, Longitudinal Invariance).

In this MTMM analyses, the critical feature is the 30 × 30 MTMM of correlations among the 15 WB-Pro factors at T1 and T2 (see Table 1). With time as the method factor, the convergent validities are the 15 test–retest correlations between matching T1 and T2 factors (correlations shaded in gray in Table 1). These are

Table 1
Test–Retest Correlations Among 15 Well-Being Factors: A Multitrait-Multimethod (MTMM) Matrix

Latent factors	CO	CT	ES	EN	ME	OP	PE	PR	RE	*SE*	VI	AC	AU	EM	PS
Correlations between Time 1 and 2 constructs: Convergent validities in diagonal (shaded), heterotrait-heteromethod correlation in off-diagonals															
Competence (CO)	.77	.33	.20	.22	.42	.24	.37	.26	.25	.31	.19	.56	.46	.15	.21
Clear thinking (CT)	.38	.75	.48	.39	.33	.34	.28	.36	.41	.36	.18	.45	.36	.15	.22
Emotional stability (ES)	.19	.41	.84	.24	.32	.33	.31	.12	.44	.30	.24	.34	.23	.04	.16
Engagement (EN)	.09	.39	.36	.78	.32	.49	.25	.35	.34	.43	.25	.31	.34	.18	.19
Meaning (ME)	.32	.33	.30	.42	.83	.59	.51	.39	.35	.38	.34	.43	.30	.15	.14
Optimism (OP)	.21	.26	.32	.54	.42	.83	.38	.35	.33	.49	.31	.34	.32	.14	.15
Positive emotions (PE)	.28	.29	.41	.42	.53	.54	.86	.37	.33	.34	.34	.50	.36	.12	.18
Positive relations (PR)	.29	.37	.36	.34	.41	.39	.44	.81	.23	.50	.17	.47	.36	.27	.26
Resilience (RE)	.25	.22	.43	.28	.32	.43	.34	.23	.78	.25	.27	.38	.28	−.07	.13
Self-esteem (SE)	.21	.23	.45	.43	.39	.42	.33	.36	.28	.80	.09	.41	.35	.18	.18
Vitality (VI)	.15	.26	.22	.37	.24	.44	.30	.16	.34	.15	.85	.24	.24	.05	.07
Acceptance (AC)	.35	.42	.41	.36	.36	.34	.29	.31	.45	.44	.18	.82	.43	.14	.35
Autonomy (AU)	.30	.35	.31	.38	.36	.35	.34	.36	.31	.31	.15	.43	.73	.11	.16
Empathy (EM)	.12	.02	.01	.13	.12	.01	.03	.23	−.04	.08	−.03	.09	.02	.82	.41
Prosocial (PS)	.14	.18	.17	.21	.27	.24	.11	.24	.21	.23	.08	.23	.24	.40	.73
Correlations among Time 1 factors (below diagonal) and among Time 2 factors (above diagonal): Heterotrait-monomethod correlations															
Competence (CO)	1.0	.61	.28	.26	.52	.31	.51	.37	.29	.42	.21	.56	.49	.15	.25
Clear thinking (CT)	.43	1.0	.47	.48	.46	.34	.44	.49	.34	.49	.26	.54	.48	.10	.29
Emotional stability (ES)	.31	.42	1.0	.35	.34	.27	.38	.30	.50	.33	.24	.50	.39	.15	.20
Engagement (EN)	.14	.41	.38	1.0	.49	.60	.40	.49	.30	.57	.34	.45	.54	.20	.29
Meaning (ME)	.49	.38	.36	.34	1.0	.57	.65	.51	.33	.44	.32	.45	.43	.09	.26
Optimism (OP)	.22	.29	.40	.54	.54	1.0	.46	.46	.37	.50	.29	.35	.47	.06	.21
Positive emotions (PE)	.42	.36	.47	.36	.59	.54	1.0	.52	.40	.42	.36	.47	.41	.09	.22
Positive relations (PR)	.35	.37	.29	.34	.47	.41	.45	1.0	.30	.57	.13	.42	.45	.28	.32
Resilience (RE)	.30	.40	.55	.44	.44	.41	.44	.27	1.0	.28	.35	.42	.31	−.05	.15
Self-esteem (SE)	.28	.41	.39	.58	.31	.46	.36	.41	.30	1.0	.06	.51	.47	.10	.26
Vitality (VI)	.20	.25	.30	.43	.34	.42	.39	.19	.31	.18	1.0	.22	.19	.01	.13
Acceptance (AC)	.45	.47	.50	.37	.50	.38	.44	.43	.49	.45	.25	1.0	.49	.15	.27
Autonomy (AU)	.39	.44	.30	.39	.42	.37	.38	.42	.34	.40	.22	.50	1.0	.11	.16
Empathy (EM)	.16	.11	.09	.15	.17	.09	.11	.29	.06	.15	.07	.22	.16	1.0	.52
Prosocial (PS)	.28	.28	.21	.28	.22	.20	.17	.31	.23	.31	.08	.35	.28	.51	1.0

Note. Table 1 is a multitrait-multimethod matrix of correlations between the 15 WB-Pro (well-being profile) factors at time 1 (T1) and time 2 (T2) for the longitudinal analysis (see Model 6, Table S2 of online supplemental materials) with time as the method factor. In the upper box, the diagonal (highlighted in gray) correlations are convergent validities (T1–T2 test–retest correlations; (.73 to .86; $M\ r = .80$); the off-diagonal values are heterotrait-heteromethod (different traits, different methods) correlations between T1 and T2 factors (−.07 to .48; $M\ r = .29$). In the lower box are heterotrait-heteromethod (different trait, different method) correlations among T1 factors (below the main diagonal, .05 to .59; $M\ r = .34$) and among T2 factors (above the main diagonal; −.05 to .65, $M\ r = .35$). Because every convergent validity is greater than all remaining (heterotrait-monomethod and heterotrait-monomethod) correlations, there is strong support for the convergent and discriminant validity of all 15 WB-Pro factors in relation to time.

consistently substantial (.73 to .86; $M\ r = .80$). This provides strong support for convergent validity in relation to time as the method factor.

In the MTMM paradigm, discriminant validity is established by comparing convergent validities to the other correlations in the MTMM matrix (see Table 1); correlations among the 15 factors among at T1 (05 to .59; $M\ r = .34$), correlations among the 15 factors at T2 (−.05 to .65, $M\ r = .35$), and correlations between T1 and T2 nonmatching factors (−.07 to .48; $M\ r = .29$). Because every convergent validity is substantially greater than all remaining correlations, there is strong support for both the convergent and discriminant validity of all 15 WB-Pro factors in relation to time. In summary, even though some of the WB-Pro factors are substantially correlated, there is clear evidence that all the factors are well differentiated based on this MTMM analysis.

Relations to Background/Demographic Characteristics

A set of 10 background/demographic characteristics were regressed on the set of WB-Pro factors (see Table S3, Section 5 of online supplemental materials). Although the resulting 150 regression coefficients are mostly modest in size, nearly half are statistically significant (due in part to the moderately large sample size). In support of our multidimensional perspective, age and gender were positively related to some WB-Pro factors and negatively related to others—results that would not be evident with unidimensional approaches to well-being. Thus, for example, males had significantly higher scores for emotional stability, resilience, vitality, and self-acceptance, but significantly lower scores for empathy, prosocial behavior, and self-esteem. However, the gender-by-age interactions demonstrated that gender difference in favor of males declined with age for emotional stability, resilience, vitality, and self-acceptance, whereas the gender difference in favor of prosocial behavior for females became larger with age. Also, in support of a multidimensional perspective, older participants had higher scores for emotional stability, clear thinking, positive emotions, resilience, and self-acceptance, but lower scores for optimism and vitality. However, there were also some quadratic effects associate with age. For example, competence increased with

age, levelled out and then declined in old age, whereas optimism initially declined with age, levelled out and then increased in old age.

The largest positive effects were associated with marital status (married = 1, not married = 0; there were significantly positive correlations for seven of the 15 factors), education (significantly positive correlations for 10 of 15 factors), and English fluency (significantly positive correlations for eight of the 15 factors, but negatively related to resiliency). The positive effects of being married did not vary as a function of age, but the advantages were slightly larger for males than females on three well-being factors (engagement, resilience, and prosocial behavior).

Relations With Significant Life Change Events

Participants indicated whether they had experienced major life events in the last 3 months. Across the 2,559 sets of responses, we classified these as negative (10.3%), neutral (5.3%), positive (6.1%), or no significant events (78%; i.e., respondents who did not list a significant life change). We contrasted the effects of not having a significant life event ("left out" category) with those for negative, neutral, or positive life events (bottom of Table S3, Section 5 of online supplemental materials).

Particularly, as most reported life events were negative, the effect of negative life events on well-being was negative for 12 of 15 WB-Pro factors, the largest being on resilience, optimism, and emotional stability (see Table S4, Section 6 of online supplemental materials). Experiencing negative life events had small positive effects on empathy and prosocial behavior (although the prosocial behavior effect was not statistically significant in relation to having no life events, but was clearly significantly different from the negative effect for most other factors). Experiencing positive life events was positively associated with well-being, although the effects were significant for only four of 15 factors (prosocial behavior, vitality, positive relations, and optimism). Not surprisingly, there were almost no differences in well-being associated with experiencing a neutral life-event change compared with not having experienced one at all (significant for only one of 15 WB-Pro factors, empathy).

Convergent and Discriminant Validity in Relation to Other Constructs

The main results from the convergent validity analysis are based on correlations in a factor analysis based on the large number of items (143) and factors (40 = 15 WB-Pro factors plus 25 covariate factors) summarized in Table 2. For the set of 375 correlations (15 WB-Pro factors × 25 external criteria), 15 are for external criteria specifically chosen to reflect a WB-Pro factor as a test of convergent validity. In support of the convergent and discriminant validity of the WB-Pro factors (see Table 2), each of the 15 convergent validities is higher than the correlation between the specific criterion of any the other WB-Pro factors. For the remaining 10 external criteria (e.g., Big-Five factors and global measures of well-being), no single WB-Pro factor was chosen to be most logically associated with each criterion. The overarching finding is that WB-Pro dimensions were generally most strongly correlated with the scales that we a prior predicted to show large associations, especially in the case of PERMA, Basic Psychological Needs, life satisfaction, and happiness. A more detailed presentation of the relations between the WB-Pro factors and scales testing convergent and divergent validity (e.g., PERMA, Basic Psychological Needs, Big-Five Personality, and other single-scale measures including depression, stress, life satisfaction, happiness, sleep, general health, and exercise) is included Section 4 of online supplemental materials.

A Profile Approach: Relations Between WB-Pro15 Factors and Selected Demographic Variables

In this section (see Table S3) we evaluate a multidimensional profile approach to the representation of the WB-Pro15 scales in relation to three demographic variables (marital status, gender, and age) and compare it to a unidimensional approach. For present purposes, we represent the unidimensional approach with responses to the Life Satisfaction measure. To be useful, the correlations between the background variable and at least some—hopefully many—of the WB-Pro15 factors must be different from the correlation with life satisfaction. Thus, for example, married respondents were nonsignificantly lower (−.10) on competence but significantly higher on life satisfaction (.23); and the difference between the two (−.10 − .23 = −.33) was significant. More generally, significantly positive difference scores (shaded dark gray) indicate that the effect of the demographic variable was significantly more positive than the corresponding effect on life satisfaction, significantly negative difference scores (shaded light gray) mean that the effect of demographic variable was more negative than for life satisfaction.

If most of the difference scores were nonsignificant, it could be argued that most of the profile of WB-Pro effects could be explained in terms of overall life satisfaction. However, a majority of the difference scores are statistically significant for all three demographics (some differences positive and others negative) and there are significant differences for all 15 WB-Pro factors. These results demonstrate good support for the multidimensional rationale underpinning WB-Pro: the profile of effects for each of these demographic variables across the WB-Pro factors cannot be explained in terms of a global measure of life satisfaction. We examine these profile differences for each of the three demographics in more detail in Section 7 of online supplemental materials.

Purportedly Unidimensional Measures of Well-Being

Of particular interest are the two widely used measures of global well-being that are purportedly unidimensional: the Diener et al. (2010) eight-item Flourishing Scale and the 14-item Warwick-Edinburgh Mental Well-being Scale (Tennant et al., 2007). Again, it is not surprising that both these global measures of well-being are substantially correlated with all WB-Pro scales (rs > .70 in bold in Table 4) and also substantially correlated with each other (r = .78). For the WEMWBS instrument the highest three correlations with WB-Pro factors are positive emotions (.82), optimism (.80), and engagement (.79), whereas for the Diener et al. (2010) instrument the largest three correlations are for meaning (.78), positive emotions (.77), and engagement (.76). For both instruments, the lowest correlations were with empathy (.25, .30) and prosocial behavior (.43, .46). Again, the patterns of correlations relating the WB-Pro items to each of these global measures of

Table 2
Correlates of Well-Being: Support for the Convergent and Discriminant Validity of 15 Well-Being Factors

Well-being correlates	Well-being profile 15 factors														
	Competence	Clear thinking	Emotional stability	Engagement	Meaning	Optimism	Positive emotion	Positive relations	Resilience	Self-esteem	Vitality	Self-acceptance	Autonomy	Empathy	Prosocial
PERMA															
Positive emotion	.68	.67	.70	.82	.81	.83	**.90**	.72	.71	.73	.72	.72	.68	.28	.43
Engagement	.69	.68	.62	**.84**	.75	.75	.80	.68	.64	.71	.66	.67	.65	.31	.46
Positive relations	.56	.58	.56	.67	.71	.69	.73	**.83**	.56	.62	.57	.59	.58	.31	.39
Meaning	.70	.65	.65	.82	**.90**	.83	.82	.70	.67	.74	.69	.70	.68	.29	.46
Accomplishment	**.72**	.69	.65	**.79**	**.81**	**.79**	**.78**	.65	.69	.71	.71	.70	.66	.24	.44
Psychological needs															
Satisfaction autonomy	.65	.63	.56	.75	.70	.68	.67	.59	.56	.62	.57	.60	**.70**	.23	.36
Satisfaction relation	.59	.58	.51	.60	.62	.59	.62	**.81**	.47	.64	.42	.57	.58	.36	.42
Satisfaction competence	**.79**	.71	.59	.67	.64	.65	.62	.57	.57	.74	.50	.67	.63	.19	.39
Frustration autonomy	−.43	−.47	−.40	−.50	−.45	−.47	−.50	−.43	−.39	−.43	−.37	−.43	**−.50**	.01	−.15
Frustration relation	−.44	−.48	−.40	−.40	−.42	−.40	−.46	**−.59**	−.35	−.50	−.25	−.43	−.41	−.07	−.20
Frustration competence	**−.58**	−.56	−.48	−.50	−.52	−.49	−.52	−.48	−.46	−.59	−.38	−.55	−.48	.02	−.19
Big Five															
Openness	.44	.39	.31	**.43**	.31	.33	.30	.30	.28	.40	.29	.37	.32	.39	**.42**
Conscientiousness	**.68**	**.65**	.45	.53	.48	.43	.44	.46	.39	.56	.39	.52	.50	.23	.42
Extraversion	.43	.38	.32	**.51**	.49	.46	**.51**	.45	.43	.48	.48	.43	.37	.26	.41
Agreeableness	.48	.45	.50	.44	.40	.37	.43	.48	.33	.50	.23	.45	.40	**.61**	**.66**
Neuroticism	−.48	−.51	**−.62**	−.48	−.47	−.48	**−.52**	−.39	**−.59**	−.48	−.44	**−.55**	−.45	.12	−.18
Single scale measures															
WEMWBS	.69	.69	.67	**.79**	**.78**	**.80**	**.82**	.72	.68	.73	.70	.72	.67	.25	.43
Flourishing	.71	.64	.62	.76	.78	.76	.77	.69	.61	.72	.62	.67	.64	.30	.46
Depression	−.45	−.53	**−.53**	−.50	−.55	−.55	**−.62**	−.48	−.51	−.51	−.50	−.54	−.45	.02	−.15
Stress	−.43	−.51	**−.60**	−.51	−.46	−.48	**−.57**	−.43	−.54	−.46	−.47	−.53	−.45	−.02	−.17
Life satisfaction	.64	.60	.60	.75	.77	**.81**	**.80**	.66	.65	.62	.69	.62	.62	.20	.35
Happiness	.51	.52	.52	.63	.63	.65	**.72**	.59	.53	.58	.52	.54	.52	.17	.35
Sleep	.07	.11	.15	.16	.18	.22	.22	.18	.17	.07	**.27**	.10	.10	.05	−.03
General health	−.37	−.35	−.36	−.42	−.40	−.46	−.41	−.31	−.39	−.32	**−.56**	−.34	−.31	−.10	−.18
Exercise	.21	.21	.25	.29	.27	.28	.25	.18	.24	.20	**.40**	.21	.20	.11	.21

Note. WEMWBS = Warwick Edinburgh Mental Well-Being Scale; WB-Pro = well-being profile. Correlations are based on a large confirmatory factor analysis model (143 items and 40 factors–15 WB-Pro15 plus 25 covariate factors). Correlations in bold are for those relations most logically and highly related to each of the WB-Pro15 factors, whereas underlined correlations are for external criteria specifically chosen a priori to reflect a WB-Pro15 factor as a test of convergent validity.

Table 3
A Profile Perspective on Relations Between the 15 Factors of WB-Pro and Selected Demographic Variables

	Married		Difference		Male		Difference		Age		Difference	
WB-Pro factors	Est.	*SE*	Est.	*SE*	Est.	*SE*	Est.	*SE*	Est.	*SE*	Est.	*SE*
Competence	−.10	.06	−.33	.06**	.08	.04*	.02	.04	.09	.06	.07	.06
Clear thinking	.12	.06*	−.11	.06*	.05	.03*	−.01	.03	.16	.03**	.15	.03**
Emotional stability	.15	.04**	−.08	.04	.12	.03**	.05	.03	.12	.04**	.10	.04**
Engagement	.25	.04**	.02	.04	−.01	.04	−.08	.04**	.00	.05	−.01	.05
Meaning	.03	.06	−.25	.06**	.07	.04*	.01	.04	.09	.06	.07	.06
Optimism	.30	.03**	.07	.03*	.00	.04	−.06	.03*	−.10	.05*	−.11	.05*
Positive emotions	.03	.05	−.26	.05**	.05	.03	−.01	.03	.13	.05*	.12	.05*
Positive relations	.17	.03**	−.06	.03*	−.03	.03	−.10	.02**	.07	.03**	.05	.02*
Resilience	.14	.04**	−.09	.04*	.13	.03**	.07	.02**	.07	.03**	.05	.02*
Self-esteem	.23	.05**	.00	.05	−.06	.03	−.12	.03**	.12	.06*	.11	.06
Vitality	.11	.03**	−.11	.03**	.16	.02**	.10	.02**	−.04	.03	−.06	.02*
Self-acceptance*	.04	.04	−.19	.04**	.12	.03**	.05	.03*	.15	.03**	.14	.03**
Autonomy*	.04	.04	−.19	.04**	.04	.03	−.03	.02	.12	.03**	.10	.03**
Empathy *	.04	.03	−.19	.04**	−.14	.03**	−.21	.03**	.00	.03	−.01	.03
Prosocial	.04	.03	−.19	.04*	−.10	.03**	−.16	.03**	.04	.03	.02	.03
Life satisfaction	.23	.02**			.06	.02**			.02	.02		

Note. WB-Pro = well-being profile. For each WB-Pro factor, we tested the effect of three demographic variables (married, male, and age) and the difference between that effect and the corresponding effect on Life Satisfaction. Thus, for example, married respondents were non-significantly lower (−.10) on competence but significantly higher on life satisfaction (.23); the difference between the two (−.10 − .23 = −.33) was significant. More generally, significantly positive difference scores (shaded dark gray) indicate that the effect of the demographic variable was significantly more positive than the corresponding effect on life satisfaction, significantly negative difference scores (shaded light gray) mean that the effect of demographic variable was more negative than for life satisfaction.
* $p < .05$. ** $p < .01$.

well-being is very similar, with a profile similarity index of .97 (i.e., the correlation between the 15 correlations relating WEMWBS to WB-Pro factors, and the corresponding 15 correlations based on Flourishing responses).

Tests of the unidimensionality of the WEMWBS and the flourishing. Both the WEMWBS and the Flourishing scales are sometimes claimed to be unidimensional measures of well-being. Although each is intended to provide a global summary score that represents global well-being, it is unclear whether these reflect unidimensional measures of a reflective well-being construct or an index of a formative measure representing different constructs. Although both reflective and formative measures can be used as global measures, the logic of their appropriate construction and derivation of psychometric properties differ substantially (for further discussion of reflective and formative measures see online supplemental material, Section 9).

We began by testing the unidimensionality of the eight-item Flourishing and 14-item WEMWBS scales, separately and in combination. Based on typical measures of goodness-of-fit (see online supplemental materials, Section 3, Table S2) and consistent with previous research, there was at least reasonable support for the unidimensionality of both the Flourishing (CFI = .96, TLI = .94, Model 8A) and WEMWBS (CFI = .93, TLI = .92, Model 8B) scales considered separately. In a CFA model with two factors based on responses to all 22 items from both instruments, the fit was somewhat poorer (CFI = .91, TLI = .89, Model 8C). In this model the Flourishing and WEMWBS factors correlate .78 with each other. The relatively poorer fit of the model with both instruments was because some items from each instrument related more strongly to some items in the other instrument than could be explained by the correlation between the two global factors. This finding is consistent with the design of each instrument to include different components of well-being that were overlapping in the two instruments, but calls into question the claims that each is a unidimensional measure.

Next, we added the WB-Pro ESEM factor structure to each of these three unidimensional models (Models 9A-C in online supplemental materials, Section 3, Table S2). Hence, in each of these models, the Flourishing items and the WEMWBS items each defined separate factors that were not allowed to cross-load on the WB-Pro factors. Thus, for example, the first WEMWBS item "I've been feeling optimistic about the future" was not allowed to load on the WB-Pro optimism factor, the factor that it was most logically related to, based on a priori classifications by the coauthors. The relation of this item to the WB-Pro optimism factor could only take place through the WEMWBS global well-being factor defined by all 14 WEMWBS items. Similarly, the first Flourishing item "My social relationships are supportive and rewarding" was not allowed to load on the WB-Pro positive relations factor and could only be related to this factor through the Flourishing Scale global well-being factor based on defined by all eight Flourishing Scale items. Again, there was at least reasonable support for the unidimensionality of the Flourishing and WEMWBS measures based on these models. For the model with one Flourishing well-being factor defined by the eight Flourishing items and 15 WB-Pro factors defined by ESEM of the 48 WB-Pro items, the fit was reasonable by typical standards of fit (CFI = .96, TLI = .92, Model 9A, Table S2 of online supplemental materials), as were the models based on the 14 WEMWBS items (CFI = .95, TLI = .92, Model 9B) and, to a lesser extent, the combination of Flourishing Scale and WEMWBS items (CFI = .91, TLI = .90). It is interesting that the fit of each of these models that included the WB-Pro items was roughly similar to the fit of the models based on the Flourishing Scale and WEMWBS items that did not include the WB-Pro items. However, in each case, the fit

Table 4
Relations Between the 15 WB-Pro Factors and Individual Items From the WEMWBS and From the Flourishing Scale: The Multidimensionality of Unidimensional Scales

WB-Pro factors		W1	W2	W3	W4	W5	W6	W7	W8	W9	W10	W11	W12	W13	W14	F1	F2	F3	F4	F5	F6	F7	F8
Competence	FL	.18	.39	.13	.18	.19	.37	.31	.23	.09	**.48**	**.45**	.13	.35	.11	.02	−.21	−.08	−.01	.36	.15	.00	.10
	PFL	.00	1.43	.00	.00	.00	1.17	1.00	.00	.00	1.73	.17	.00	.00	.00	.00	.00	.00	.00	1.67	.00	.00	.30
Clear think	FL	−.07	.08	.03	−.03	−.07	.13	**.59**	−.07	−.12	.00	.36	.00	.13	−.11	−.12	−.02	.20	.06	.32	.11	−.11	.04
	PFL	.00	.00	.00	.00	.00	.66	3.00	.00	.00	.00	1.67	.00	.00	.00	.00	.00	.00	.00	1.17	.00	.00	.00
Emotional stability	FL	.08	.05	.29	.08	.06	.21	.32	.10	.01	.05	.11	−.02	−.08	.13	.02	−.02	−.04	−.08	−.08	.07	.10	.07
	PFL	.00	.00	1.67	.00	.00	.30	.00	.00	.00	.00	.00	.33	.00	.33	.00	.00	.00	.00	.00	.00	.00	.00
Engagement	FL	.12	.24	.33	.37	.39	.11	.25	.17	.08	.20	.14	−.03	**.53**	.24	.15	.19	**.61**	.19	.14	−.09	−.16	−.16
	PFL	.00	.30	.67	1.43	.67	.00	.00	.00	.00	.10	.00	.00	2.83	.00	.33	.00	2.67	.00	.00	.00	.00	.00
Meaning	FL	.09	.22	−.07	−.05	−.06	.02	.01	.04	−.10	.13	.02	−.03	−.07	−.05	**.76**	.09	.17	.14	.15	.21	.33	.22
	PFL	.33	1.57	.00	.00	.00	.00	.00	.17	.00	.00	.00	.00	.33	.00	3.00	.00	.67	.17	.33	2.00	.67	.00
Optimism	FL	**.73**	.04	.04	.04	.04	−.01	.00	.11	−.04	.00	−.02	.01	.23	.11	.05	−.12	−.07	.03	.05	.06	**.63**	−.09
	PFL	3.33	.00	.00	.00	.00	.00	.00	.17	.00	.00	.00	.00	.33	.00	.00	.00	.00	.00	.00	.00	3.00	.00
Positive emotions	FL	.12	.20	.36	.14	.07	.20	.09	.39	.20	.20	.06	.18	.10	**.50**	.13	.07	.14	−.05	−.22	.08	−.02	−.08
	PFL	.33	.00	1.33	.00	.67	.00	.00	1.00	.33	.47	.00	1.33	.17	3.33	.33	.00	.00	.00	.00	.17	.33	.10
Positive relations	FL	.02	.10	.04	.31	−.02	.01	.05	.07	**.74**	.02	−.03	**.83**	−.03	.12	.14	**.82**	.13	.33	.12	.15	.07	**.42**
	PFL	.00	.00	.00	1.57	.00	.00	.00	.00	2.67	.00	.00	2.33	.00	.00	.00	3.67	.00	1.67	.00	.00	.00	1.87
Resilience	FL	−.06	.06	.04	.03	−.05	.19	−.11	−.06	−.05	.09	.02	−.08	.07	−.03	−.08	−.06	−.05	−.01	.01	−.03	−.13	.03
	PFL	.00	.00	.00	.00	.00	1.51	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
Self-esteem	FL	.00	.16	.01	−.08	−.04	.23	.03	.22	−.07	.31	.19	−.02	−.05	.18	.25	.02	.14	.25	**.61**	**.72**	.35	**.47**
	PFL	.00	.70	.00	.00	.00	.59	.00	2.67	.00	1.70	.00	.00	.00	.00	.33	.00	.00	.00	.83	1.83	.00	1.40
Vitality	FL	−.08	.02	.04	−.06	**.58**	.01	−.03	.07	−.03	.03	−.11	−.07	.03	.07	.10	.11	*.21*	.01	.02	.03	.05	.14
	PFL	.00	.00	.00	.00	2.67	.00	.00	.00	.00	.00	.00	.00	.00	.33	−.03	−.05	−.02	−.15	−.07	−.07	.00	.06
Self-accept	FL	−.05	−.07	−.02	−.10	−.06	−.10	−.05	.09	.00	.11	−.07	−.08	−.10	−.06	−.09	−.04	−.01	−.02	−.03	−.05	−.07	−.13
	PFL	.00	.00	.33	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
Autonomy	FL	−.01	−.12	.05	−.04	−.02	.06	.05	.03	−.01	.01	.31	−.02	−.01	.06	−.05	−.10	−.01	−.16	−.02	−.01	−.04	.07
	PFL	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	2.17	.00	.33	.00	.00	.00	.33	.00	.00	.00	.00	.00
Empathy	FL	.02	.04	.08	.33	.18	.01	−.04	−.07	.19	−.03	−.09	.08	.13	.10	−.08	.08	−.06	−.02	−.06	−.04	−.09	−.03
	PFL	.00	.00	.00	.53	.00	.00	.00	.00	.50	.00	.00	.00	.00	.00	.00	.00	.00	.33	.00	.00	.00	.00
Prosocial	FL	.02	.11	.02	.22	.04	.04	.03	.10	.22	.00	.05	.13	.16	.05	.04	.11	.07	**.44**	.03	.03	−.01	.09
	PFL	.00	.00	.00	.47	.00	.00	.00	.00	.50	.00	.00	.00	.00	.00	.00	.33	.00	1.83	.00	.00	.00	.33

Note. WEMWBS = Warwick Edinburgh Mental Well-Being Scale; WB-Pro = well-being profile. FL = actual factor loadings of the 14 WEMWBS and eight Flourishing items on each of the WB-Pro factors (those over .20 are in bold and the highest for each item is highlighted in gray). PFL = predicted factor loading, summated values from four co-authors as to which WB-Pro factors the 22 items would be most correlated with (highest possible score is 4). Correlation = the correlation between FL and PFL values for each the 15 WB-Pro factors (last column) and for each of the 14 + 8 =24 items (the last row); the correlation across all 330 (22 items × 15 factors) sets of FLs and EFLs is .81. 14 WEMWBS (W) items: W1, I've been feeling optimistic about the future. W2, I've been feeling useful. W3, I've been feeling relaxed. W4, I've been feeling interested in other people. W5, I've had energy to spare. W6, I've been dealing with problems well. W7, I've been thinking clearly. W8, I've been feeling good about myself. W9, I've been feeling close to other people. W10, I've been feeling confident. W11, I've been able to make up my own mind about things. W12, I've been feeling loved. W13, I've been interested in new things. W14, I've been feeling cheerful. 8 Flourishing (F) items: F1, I lead a purposeful and meaningful life. F2, My social relationships are supportive and rewarding. F3, I am engaged and interested in my daily activities. F4, I actively contribute to the happiness and well-being of others. F5, I am competent and capable in the activities that are important to me. F6, I am a good person and live a good life. F7, I am optimistic about my future. F8, People respect me.

of these three models was noticeably poorer than the fit of models considered earlier based on only the WB-Pro items.

Absorption of WEMWBS and Flourishing items into WB-Pro. Our final set of models (10A–10C, Table S2, online supplemental materials) was specifically designed to test our a priori hypothesis that the Flourishing Scale and WEMWBS are multidimensional instruments. More specifically, we tested extended target ESEM models in which the Flourishing Scale and WEMWBS items were fully "absorbed" into the WB-Pro factors, completely eliminating the Flourishing Scale and WEMWBS global factors. As with the WB-Pro, each of the WEMWBS and Flourishing Scale items was designated a priori as a "target item" to the most closely related to a WB-Pro factor or as a "nontarget item" based on the coauthors' a priori classifications of items into WB-Pro factor factors (see Table 4; also see Mplus syntax, Section 11). Thus, all three models (48 WB-Pro items with eight Flourishing items, Model 10A, Table S2 of online supplemental materials; with 14 WEMWBS items, Model 10B; or with 22 items from both Flourishing Scale and WEMWBS items, Model 10C) posited only the 15 WB-Pro factors, such that all the additional WEMWBS and Flourishing items were absorbed into the WB-Pro factors. The critical evaluations of these models are: first, do these "absorption models" fit systematically better than those already considered in which the WEMWBS and Flourishing Scale items define separate factors; and second, do the WEMWBS and Flourishing Scale items load on the WB-Pro factors with which they are most logically related (based on a priori classifications by the coauthors)?

For each of the three models, the absorption models fit substantially better than the corresponding models hypothesizing separate global well-being factors defined by the WEMWBS or Flourishing Scale items (see Models 8, 9 and 10 in Table S2 of online supplemental materials). Of particular relevance are the models based on the 48 WB-Pro items, 14 WEMWBS items, and eight Flourishing Scale items. As already noted, the model based on 17 factors (15 WB-Pro, 1 WEMWBS, 1 Flourishing) provided a marginal fit to the data (RMSEA = 0.066, CFI = .91, TLI = .89, Model 9C). In contrast the corresponding absorption model based on only 15 factors provided a much better fit to the data (RMSEA = 0.022, CFI = .98, TLI = .97, Model 10C). A similar pattern of results is evident for corresponding models that considered the WEMWBS and Flourishing Scale items separately.

In Table 4, we present the factor loadings for the 14 WEMWBS and eight Flourishing Scale items on the WB-Pro factors. (To facilitate presentation, those over .20 are shaded in gray and those greater than .35 are bolded.) Also presented are the a priori predictions regarding the 22 items of WB-Pro factors that would be most correlated (the highest possible score is 4). Thus, for example, the first WEMWBS item "I've been feeling optimistic about the future" was classified as fitting into the optimism WB-Pro factor (classification score = 3.33) and the item loaded .73 on this factor. Similarly, the first Flourishing Scale item, "I lead a purposeful and meaningful life," was classified into the meaning WB-Pro factor (classification score = 3.00) and the item loaded .76 on this factor. All 22 Flourishing and WEMWBS items load at least moderately or substantially on one or more WB-Pro factors. The pattern of factor loadings of the 22 items on the 15 factors is closely aligned to the a priori classifications. Thus, the profile similarity index, the correlation between the set of 330 (22 items × 15 factors) factor loadings and the corresponding set of 330 classifications, is .81.

Although there is good support for the a priori classification of the Flourishing Scale and WEMWBS items into different WB-Pro factors, the coverage of items across the 15 WB-Pro factors is not uniform. WEMWBS items are substantially related to competence, engagement, positive emotions, and self-esteem with very little representation of meaning, resilience, self-acceptance, autonomy, empathy, or prosocial behavior. The Diener et al. (2010) Flourishing Scale items are most strongly represented in the WB-Pro self-esteem factor (six of eight items load substantially on this factor—those shaded in Table 3). However, seven of the eight items load most substantially on a single WB-Pro factor (the item "People respect me" loads substantially on both WB-Pro factors positive relations and self-esteem). Nevertheless, the Diener et al. (2010) instrument has relatively little coverage of emotional stability, positive emotions, resilience, vitality, self-acceptance, autonomy, or empathy. This analysis of the Flourishing and WEMWBS instruments also highlights differences, as well as similarities, in the components of well-being covered by each of these instruments. Thus, for example, the Flourishing—but not the WEMWBS—instrument has items specifically targeted to reflect engagement, meaning, self-esteem, and prosocial behavior. In contrast, the WEMWBS—but not the Flourishing—instrument has items specifically targeted to reflect clear thinking, competence, vitality, and, perhaps, autonomy.

Development of a Short-Form of WB-Pro

Recognizing the usefulness of a much briefer measure of well-being than the WB-Pro (with 48 items), we sought to develop an explicitly formative measure of well-being based on a single global measure rather than a multidimensional profile of distinct factors. It is not our intention to develop a relatively unidimensional (reflective) measure of well-being based on items that are highly internally consistent, rather we aim here to develop a formative index measure of well-being that most appropriately encompasses the range of content covered by all WB-Pro factors. This has been accomplished with an innovative machine-learning approach based on selections of subsets of items that explain the maximum amount of variance in the total set of 48 WB-Pro items (for further discussion see online supplemental materials, Section 8).

For present purposes, we constructed two global scales from the 48 WB-Pro items, one based on the best 15 items subject to the constraint that one item was included from each of the 15 WB-Pro factors, and one based on the best five items. In evaluating these short forms, we focused on two issues: First, we examined how related the global scales are to each of the 25 external criteria (online supplemental materials, Table S5); second, we examined whether the global scales are as highly related to any of the criteria as the highest correlating WB-Pro factor. This second step is a test of the multidimensionality of well-being, because support for the multidimensional perspective requires that the global scores are less correlated with the external criteria than the highest correlating WB-Pro factor. The pattern and even the size of correlations between the set of 25 external criteria and each of the two global scores is similar (online supplemental materials, Table S5). Although the 15-item global scale does marginally better than the five-item global scale, the difference

is minimal, and the five-item scale actually does better for some of the criteria (for further discussion see online supplemental material, Section 10).

Discussion

We aimed to develop a robust multidimensional measure of well-being, based on the dimensions identified in the systematic approach used by Huppert and So (2013), and drawing on an extensive review of the subjective well-being literature. Defining psychological well-being as the opposite of psychological ill-being (the common mental disorders of anxiety and depression), our approach is among the most systematic attempts to-date to define and measure well-being, and resulted in a novel, theory-based measure of subjective well-being.

Contrasting Purportedly Unidimensional and Multidimensional Measures of Well-Being

The multidimensional approach we have used to measure well-being is in sharp contrast to unidimensional approaches. In one of the unidimensional approaches, well-being is inferred from responses to a single item (e.g., "happiness" or "life satisfaction") or a tightly worded set of items designed to measure a narrowly defined construct. Such an approach is truly unidimensional, highly parsimonious and expedient. However, this approach provides a very narrowly defined measure of well-being and does not provide useful information about the profile of different components that make up well-being. As a result, it cannot provide practical guidance with respect to policy, or the choice of specific interventions to improve well-being components (Huppert & Ruggeri, 2018).

In a second unidimensional approach, illustrated by the widely used Flourishing Scale (Diener et al., 2010) and WEMWBS (Tennant et al., 2007), well-being is based on responses to a set of items implicitly designed to cover the well-being construct in greater breadth. Clearly, this approach results in a more broadly defined measure of well-being; however, because well-being is still represented by a single score, it does not provide useful information about the profile of different components that make up well-being, or even the components used to construct the measure. The range of well-being content sampled by these measures has been compromised to maximize internal consistency (i.e., items were dropped that were not internally consistent). Furthermore, although purportedly unidimensional, the explicit logic of the design of these instruments is multidimensional—covering a range of different components of well-being. Additionally, the rationale for what components of well-being are included or not in the implicit definition of well-being based on a single total score, is often unclear.

Our study demonstrated the limitations of treating purportedly multidimensional scales as unidimensional. While our study focused on the WEMWBS and Flourishing scales as illustrations of this, our findings may have implications for other widely used well-being scales (e.g., the Comprehensive Inventory of Thriving; Su et al., 2014).

A Multidimensional Perspective: The WB-Pro Factor Structure

We found very strong empirical support for the multidimensional factor structure of the WB-Pro instrument. First, the factor structure of the 48-item instrument, using both ESEM and CFA, provided an extremely good fit to the data and support for the a priori 15-factor structure. Although the CFA structure was more parsimonious (i.e., required fewer parameter estimates), the ESEM structure fit the data somewhat better and resulted in more differentiates, distinct (i.e., less correlated) factors. Whereas we prefer the ESEM factor structure, results based on both are very similar and both provide strong support for the a priori WB-Pro factor structure.

Second, we found very strong support for the invariance (metric and scalar as well as configural) of the factor structure over levels of education, age, gender, and time for our nationally representative sample of U.S. adults. This support for invariance demonstrated the robustness of the WB-Pro factor structure and the appropriateness of comparing scores across these different demographic groups.

Relations with demographic variables. In support for our multidimensional perspective (but also substantively relevant), we found distinct patterns of relations between the 15 WB-Pro factors and 10 demographic variables (including age, gender, education, and marital status). Although this pattern of relations is substantively interesting in its own right, the overarching insight is that this pattern of relations could not be represented with a single global measure of well-being, thereby underscoring the value of taking a multidimensional approach to the study of well-being.

Life change events. To examine how the WB-Pro dimensions relate differentially to the experience of significant life-events, we asked participants to describe significant life events experienced during the previous 3 months. These were classified as positive, negative, neutral, or not having occurred. Consistent with our finding of differential relations between the WB-Pro dimensions and demographic variables, we found that having negative life experiences were negatively associated with most of the WB-Pro well-being factors (especially optimism, resilience, and emotional stability), whereas negative life events were unrelated to competence and positively related to empathy and prosocial behavior. This suggests that life set-backs may inhibit more hedonic aspects of well-being, but strengthen eudemonic aspects, such as connecting with and assisting others. There were fewer responses for positive and neutral events, limiting the conclusions we can draw from these, but the overall picture suggests a varied pattern of associations between the 15 WB-Pro dimensions and significant life events—positive, negative, and neutral—further reinforcing the value of a multidimensional approach to the study of well-being.

A Multidimensional Perspective: Support for Convergent and Discriminant Validity

MTMM analysis in relation to time. The MTMM paradigm (Campbell & Fiske, 1959) is specifically designed to test for convergent and discriminant validity. Here, we operationalized the multiple methods in relation to time such that convergent validity was based on test–retest correlations. The MTMM design in rela-

tion to time provides a strong basis for evaluating test–retest stability and discriminant validity. Across the WB-Pro factors, we found strong support for test–retest stability. However, stability over time would normally be classified as falling near the reliability end of the reliability-validity continuum. For that reason, we now turn to validity criteria that provide a stronger test of external validity.

Relations with selected set of 25 external validity criteria. We evaluated the construct (convergent and discriminant) validity of the WB-Pro factors in relation to a purposively selected set of 25 external validity criteria. Across all WB-Pro factors and 25 criteria (375 correlations in Table 2), 15 criteria were selected a priori as closely matching different WB-Pro factors—convergent validities. In support of convergent and discriminant validity, each of these 15 convergent validities was statistically significant and larger than correlations with any other WB-Pro factor.

The WB-Pro factors overlap substantially with four of five PERMA factors (all but accomplishment) and the SDT's six basic psychological needs factors (particularly the needs satisfaction measures but also, to a lesser extent the needs frustration measures). Consistent with this a priori matching, all 10 of these convergent validities (4 PERMA and 6 Psychological Needs) are substantial (rs = .70 to .90 for the seven positively oriented external criteria, and −.50 to −.58 for the three negatively oriented need frustration factors). In support of discriminant validity, all 10 convergent validities are higher than correlations with any of the other WB-Pro factor. Also, given the focus of the WB-Pro on positive well-being, it is not surprising that the needs frustration factors are less correlated with WB-Pro factors than the corresponding need satisfaction factors.

Several of the single-scale measures (i.e., depression, stress, happiness, general health, and exercise) were specifically selected a priori as being most closely associated with specific WB-Pro factors (positive emotions, emotional stability, positive emotions, vitality, and vitality, respectively). Although the convergent validities relating these two sets of measures tend to be smaller, the pattern of results again supports both the convergent and discriminant validity of the WB-Pro responses. Other external validity criteria were not specifically linked to particular WB-Pro factors, but the results demonstrate that there was a clear, logical pattern of relations in support of convergent and discriminant validity. Thus, for example, the Big-Five agreeableness factor was most highly correlated with WB-Pro prosocial behavior and empathy, whereas the Big-Five conscientiousness factor was most highly correlated with WB-Pro competence and clear thinking. In summary, relations between the WB-Pro factors and the set of 25 external criteria provide strong support for both the convergent and discriminant validity of WB-Pro responses.

Global Measures of Well-Being: Reflective Versus Formative Measures

Are the WEMWBS and Flourishing Scale measures really global unidimensional measures? The well-being research literature is divided on the use of multidimensional and global approaches to well-being. The problem with the multidimensional approach is that there is no consensus on the factors that should be included in the instrument, nor even a conceptual or theoretical framework to guide these decisions. Among those who opt for a global measure, some use a very narrow unidimensional approach based on a single item or a small number of highly internally consistent items (e.g., measures of life satisfaction or happiness considered here). However, others define a global measure of well-being as a broad formative index based on diverse components of well-being (e.g., Flourishing Scale and WEMWBS measures considered here). Furthermore, there appears to be confusion in the use of global unidimensional and global formative measures. Thus, for example, the Flourishing Scale and WEMWBS measures are sometimes claimed to be unidimensional. However, to the extent that these are designed to be broad, formative indices of well-being, unidimensionality (and high internal consistency) is an inappropriate criterion and antithetical to the rationale underlying formative indices. Indeed, the logic of formative indices is to specifically choose a number of highly differentiated indicators that cover as broad a range of relevant content as possible. Hence, if unidimensionality was used to select items in these measures, it would detract from their usefulness. This confusion is evident, for example, with the WEMWBS instrument in which a short seven-item version (Stewart-Brown et al., 2009) of the longer 14-item (Tennant et al., 2007) version of the instrument was constructed based on strategies to maximize internal consistency. In contrast, a good formative measure should be designed to minimize redundancy (and internal consistency) among indicators to cover as much of the relevant content with as few items as possible.

In the present investigation, we demonstrated apparent problems with the widely used Flourishing Scale and WEMWBS measures, as global, unidimensional measures of well-being. In particular, we began by showing marginal support for unidimensionality of both instruments in relation to goodness-of-fit criteria. However, we also demonstrated that both of the instruments cover a broad selection of WB-Pro factors, which clearly supports a multidimensional perspective. We then considered factor analyses that combined the items from these instruments with the 48 items from the WB-Pro instrument. In these analyses, models that constrained the WEMWBS or the Flourishing Scale items to be a single factor fit worse than models that allowed WEMWBS and/or Flourishing Scale items to load on the WB-Pro factors selected a priori as being most closely associated with each of the items. In this way, the WEMWBS and/or Flourishing Scale items were absorbed into the WB-Pro factor structure. Not only did these absorbed models provide a much-improved fit to the data, but the empirical factor loadings (relating the WEMWBS and/or Flourishing Scale items to the WB-Pro factors) closely matched our a priori predictions based on the content of each of the Flourishing Scale and/or WEMWBS items. These results apparently resolve at least some of the confusion about the inappropriate role of unidimensionality in formative index measures that are intended to sample from diverse content.

Global formative measures based on the WB-Pro. In addition to our 48-item WB-Pro instrument, we specifically constructed two global, formative-index measures based on subsets of the 48 WB-Pro items using an innovative machine-learning approach. The rationale for this approach was to sample diverse content from within the item pool to maximize variance explained in the 48 items with a minimal subset of items. Using this method, we constructed two global measures, one consisting of 15 items and another five items. We then evaluated these global scales in relation to our set of 25 external criteria (that were not used in the selection of items in these global scales). Consistent with the

proposed usefulness of the global measures, both global measures were substantially related to each of the 25 external criteria. Not surprisingly, the 15-item scale did marginally better than the five-item scale in terms of predicting most—but not all—of the 25 criteria. However, consistent with our multidimensional perspective, neither of the global measures was as highly correlated with the most logically related WB-Pro factor.

What Are the Implications of These Results in Terms of Future Application of the WB-Pro?

Clearly, if researchers are interested in the multidimensionality of well-being and can justify the inclusion of all 48 WB-Pro items, then we recommend the use of the entire instrument. Because of our focus on a multidimensional perspective, we are loath to recommend the use of a global scale instead of the full WB-Pro instrument. However, if researchers can only justify the use of a relatively small number of items, we recommend the use of a broadly defined formative index of well-being such as our 15-item measure (WB-Pro15) that incorporates all 15 WB-Pro dimensions, or the very brief five-item measure (WB-Pro5). Nevertheless, because the 15- and five-item global WB-Pro measures are new, it is important that further research juxtaposes their usefulness in relation to more widely used measures of global well-being, such as the WEMWBS (Tennant et al., 2007) and the Flourishing Scale (Diener et al., 2010) and the brief inventory of thriving (BIT; Su, Tay, & Diener, 2014). However, it is also important that such further research explores in greater detail the apparent confusion between global unidimensional and global formative measures of well-being.

Using WB-Pro to Map the Content of Alternative Measures of Well-Being

The analyses we have reported, particularly those based on the items in the WEMWBS (Tennant et al., 2007) and Flourishing Scale (Diener et al., 2010), also provide a heuristic demonstration that the WB-Pro provides a suitable base to map the content of alternative measures of well-being and related constructs. Extending this mapping metaphor, Marsh, Hau, Artelt, Baumert, and Peschar (2006) suggested that a broadly based multidimensional instrument can provide the latitude and longitude for mapping the content of different measures widely used within a specific research literature. Also, related to this issue are jingle-jangle fallacies (Marsh, Pekrun, et al., 2019) where two factors with the same name do not necessarily measure the same content (jingle fallacy), or two factors with different labels might measure the same content (jangle fallacy). Using this approach, we were able to map the content covered by WEMWBS and the Flourishing Scale. In particular, by mapping the WEMWBS and Flourishing items onto the WB-Pro factors, we not only demonstrated that neither of these global measures should be considered unidimensional, but also demonstrated that the content covered by each was not the same. Using the same approach, we were also able to map the content of the Big-five, PERMA, and SDT factors onto WB-Pro factors. This use of the WB-Pro instrument to map the content of alternative measures seems particularly relevant in an area like well-being measurement that Diener and Seligman (2004, p. 2) described as "haphazard, with different studies assessing different concepts in different ways."

Limitations, Conclusions, and Directions for Further Research

WB-Pro is a 15-factor comprehensive measure of subjective well-being, based on a sound conceptual model and strong empirical support. The 15 factors showed good reliability, test–retest correlation, convergent/discriminant validity in relation to stability over time and relevant psychological measures, and a good a priori fit to the data that was invariant over time, education, gender, and age. We note, however, that our systematic approach to selection of factors might have excluded some potentially important factors. Furthermore, substantial correlations among some factors might detract from their discriminant validity. Also, the sample of respondents was U.S. adults so that further research is needed to test the generalizability to other ages and nationalities.

There are a number of ways in which this new measure, the WB-Pro, can advance research and policies related to well-being. First, it facilitates an understanding of how multidimensional profiles vary between individuals and groups. Second, it will allow us to track change and explore how different dimensions change over time, as a result of societal/cultural change, or following an intervention program. Further, having identified differences between individuals or groups, we can design policies and programs to enhance well-being that are tailored to the profiles we observe (Huppert & Ruggeri, 2018). For example, if one group obtains high scores on measures such as engagement, competence, or meaning but low scores on other measures such as autonomy, emotional stability, or self-acceptance, we can tailor programs to focus on those dimensions where they show weakness rather than on dimensions where they show strength. On the other hand, we acknowledge that not all research studies are able to include a 48-item measure of well-being, and we have accordingly used a machine-learning approach (genetic algorithm methodology) to develop two shorter forms based on these 48 items: the WB-Pro5 and the WB-Pro15. Both the 5- and 15-item measures are robust short measures of well-being that represent the range of WB-Pro factors, and can each be used as a single total score. Similar to our claims in relation to the WEMWBS and Flourishing Scale, we stress that these short measures based on the WB-Pro instrument should be considered formative rather than reflective measures. Our short global measures can be used where there are strict limitations on survey length, but it is worth recognizing that the full WB-Pro instrument provides more reliable and robust, multidimensional representations of well-being (to download short and long versions see https://ippe.acu.edu.au/research/research-instruments/wb-pro for long and short versions of the instrument).

References

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.

Asparouhov, T., & Muthén, B. (2009). Exploratory structural equation modeling. *Structural Equation Modeling, 16,* 397–438. http://dx.doi.org/10.1080/10705510903008204

Baron-Cohen, S. (2011). *Zero degrees of empathy: A new theory of human cruelty*. London, UK: Allen Lane.

Basarkod, G., Sahdra, B., & Ciarrochi, J. (2018). Body image-acceptance and action questionnaire-5: An abbreviation using genetic algorithms. *Behavior Therapy, 49,* 388–402. http://dx.doi.org/10.1016/j.beth.2017.09.006

Batson, C. D. (1991). *The altruism question: Toward a social-psychological answer*. Hillsdale, NJ: Erlbaum, Inc.

Bollen, K. A., & Lennox, R. (1991). Conventional wisdom on measurement: A structural equation perspective. *Psychological Bulletin, 110,* 305–314. http://dx.doi.org/10.1037/0033-2909.110.2.305

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

Butler, J., & Kern, M. L. (2016). The PERMA-Profiler: A brief multidimensional measure of flourishing. *International Journal of Wellbeing, 6,* 1–48. http://dx.doi.org/10.5502/ijw.v6i3.526

Caligor, E., Levy, K. N., & Yeomans, F. E. (2015). Narcissistic personality disorder: Diagnostic and clinical challenges. *The American Journal of Psychiatry, 172,* 415–422. http://dx.doi.org/10.1176/appi.ajp.2014.14060723

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56,* 81–105. http://dx.doi.org/10.1037/h0046016

Campbell, D. T., & O'Connell, E. J. (1967). Methods factors in multitrait-multimethod matrices: Multiplicative rather than additive? *Multivariate Behavioral Research, 2,* 409–426. http://dx.doi.org/10.1207/s15327906mbr0204_1

Cantril, H. (1965). *The pattern of human concerns*. New Brunswick, NJ: Rutgers University Press.

Cappelleri, J. C., Bushmakin, A. G., McDermott, A. M., Sadosky, A. B., Petrie, C. D., & Martin, S. (2009). Psychometric properties of a single-item scale to assess sleep quality among individuals with fibromyalgia. *Health and Quality of Life Outcomes, 7,* 54. http://dx.doi.org/10.1186/1477-7525-7-54

Chen, B., Vansteenkiste, M., Beyers, W., Boone, L., Deci, E. L., Van der Kaap-Deeder, J., . . . Verstuyf, J. (2015). Basic psychological need satisfaction, need frustration, and need strength across four cultures. *Motivation and Emotion, 39,* 216–236. http://dx.doi.org/10.1007/s11031-014-9450-1

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling, 14,* 464–504. http://dx.doi.org/10.1080/10705510701301834

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

Clark, A. E., Flèche, A., Layard, S., Powdthavee, R., & Ward, N. (2018). *The origins of happiness: The science of well-being over the life course*. Princeton, NJ: Princeton University Press. http://dx.doi.org/10.2307/j.ctvc77dgq

Cummins, R. (1996). The domains of life satisfaction: An attempt to order chaos. *Social Indicators Research, 38,* 303–328. http://dx.doi.org/10.1007/BF00292050

Davis, M. H. (1994). *Empathy: A social psychological approach*. Madison, WI: Brown & Benchmark.

Davis, M. H., & Oathout, H. A. (1987). Maintenance of satisfaction in romantic relationships: Empathy and relational competence. *Journal of Personality and Social Psychology, 53,* 397–410. http://dx.doi.org/10.1037/0022-3514.53.2.397

Dicke, T., Marsh, H. W., Riley, P., Parker, P. D., Guo, J., & Horwood, M. (2018). Validating the Copenhagen Psychosocial Questionnaire (COPSOQ-II) using set-ESEM: Identifying psychosocial risk factors in a sample of school principals. *Frontiers in Psychology, 9,* 584. http://dx.doi.org/10.3389/fpsyg.2018.00584

Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The satisfaction with life scale. *Journal of Personality Assessment, 49,* 71–75. http://dx.doi.org/10.1207/s15327752jpa4901_13

Diener, E., & Seligman, M. E. (2004). Beyond money: Toward an economy of well-being. *Psychological Science in the Public Interest, 5,* 1–31. http://dx.doi.org/10.1111/j.0963-7214.2004.00501001.x

Diener, E., Suh, E. M., Lucas, R. E., & Smith, H. L. (1999). Subjective well-being: Three decades of progress. *Psychological Bulletin, 125,* 276–302. http://dx.doi.org/10.1037/0033-2909.125.2.276

Diener, E., Wirtz, D., Tov, W., Kim-Prieto, C., Choi, D., Oishi, S., & Biswas-Diener, R. (2010). New well-being measures: Short scales to assess flourishing and positive and negative feelings. *Social Indicators Research, 97,* 143–156. http://dx.doi.org/10.1007/s11205-009-9493-y

Edwards, J. R., & Bagozzi, R. P. (2000). On the nature and direction of relationships between constructs and measures. *Psychological Methods, 5,* 155–174. http://dx.doi.org/10.1037/1082-989X.5.2.155

Eisenbarth, H., Lilienfeld, S. O., & Yarkoni, T. (2015). Using a genetic algorithm to abbreviate the Psychopathic Personality Inventory-Revised (PPI-R). *Psychological Assessment, 27,* 194–202. http://dx.doi.org/10.1037/pas0000032

Eisenberg, N., Fabes, R. A., & Spinrad, T. L. (2007). Prosocial development. In W. Damon, R. M. Lerner, & N. Eisenberg (Eds.), *Handbook of child psychology: Vol. 3. Social, emotional, and personality development* (pp. 646–718). New York, NY: Wiley.

Forsyth, J. P., Parker, J. D., & Finlay, C. G. (2003). Anxiety sensitivity, controllability, and experiential avoidance and their relation to drug of choice and addiction severity in a residential sample of substance-abusing veterans. *Addictive Behaviors, 28,* 851–870. http://dx.doi.org/10.1016/S0306-4603(02)00216-2

Fredrickson, B. (2009). *Positivity*. New York, NY: Three Rivers Press.

Gogol, K., Brunner, M., Goetz, T., Martin, R., Ugen, S., Keller, U., . . . Preckel, F. (2014). "My questionnaire is too long!" The assessments of motivational-affective constructs with three-item and single-item measures. *Contemporary Educational Psychology, 39,* 188–205. http://dx.doi.org/10.1016/j.cedpsych.2014.04.002

Hayes, S. C., Strosahl, K., & Wilson, K. G. (1999). *Acceptance and commitment therapy: An experiential approach to behavior change*. New York, NY: Guilford Press.

Hayes, S. C., Strosahl, K. D., & Wilson, K. G. (2012). *Acceptance and commitment therapy: The process and practice of mindful change* (2nd ed.). New York, NY: Guilford Press.

Hayes, S. C., Wilson, K. G., Gifford, E. V., Bissett, R., Piasecki, M., Batten, S. V., . . . Gregg, J. (2004). A preliminary trial of Twelve-Step Facilitation and Acceptance and Commitment Therapy with polysubstance-abusing methadone-maintained opiate addicts. *Behavior Therapy, 35,* 667–688. http://dx.doi.org/10.1016/S0005-7894(04)80014-5

Henry, C. S., Sager, D. W., & Plunkett, S. W. (1996). Adolescent's perceptions of family system characteristics, parent-adolescent dyadic behaviors, adolescent qualities, and adolescent empathy. *Family Relations, 45,* 283. http://dx.doi.org/10.2307/585500

Holland, J. (1975). *Adaption in natural and artificial systems*. Ann Arbor, MI: The University of Michigan Press.

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

Huppert, F. A. (2014). The state of well-being science: Concepts, measures, interventions and policies. In F. Huppert & C. Cooper (Eds.), *Wellbeing: A complete reference guide, Volume VI, Interventions and policies to enhance wellbeing* (pp. 1–49). Chichester, England: Wiley-Blackwell. Retrieved from https://www.safaribooksonline.com/library/view/wellbeing-a-complete/9781118716168/9781118716168c01.xhtml

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,* 301–315.

Huppert, F., Marks, N., Michaelson, J., Vázquez, C., & Vittersø, J. (2013). *European social survey round 6 module on personal and social wellbeing—Final module in template*. London, UK: European Social Survey. Retrieved from http://www.europeansocialsurvey.org/docs/round6/questionnaire/ESS6_final_personal_and_social_well_being_module_template.pdf

Huppert, F. A., & Ruggeri, K. (2018). Policy challenges: Well-being as a priority in public mental health. In D. Bhugra, K. Bhul, S. Wong, & S. Gillman (Eds.), *Oxford textbook of public mental health* (pp. 131–140). Oxford, England: Oxford University Press.

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

Jahoda, M. (1958). *Current concepts of positive mental health*. New York, NY: Basic Books. http://dx.doi.org/10.1037/11258-000

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning*. New York, NY: Springer. http://dx.doi.org/10.1007/978-1-4614-7138-7

Kahneman, D., Diener, E., & Schwarz, N. (Eds.). (1999). *Well-being: Foundations of hedonic psychology*. New York, NY: Russell Sage Foundation.

Keyes, C. L., & Waterman, M. B. (2003). Dimensions of well-being and mental health in adulthood. In M. Bornstein (Ed.), *Crosscurrents in contemporary psychology* (pp. 477–497). Mahwah, NJ: Erlbaum.

Kotov, R., Krueger, R. F., Watson, D., Achenbach, T. M., Althoff, R. R., Bagby, R. M., . . . Zimmerman, M. (2017). The Hierarchical Taxonomy of Psychopathology (HiTOP): A dimensional alternative to traditional nosologies. *Journal of Abnormal Psychology, 126,* 454–477. http://dx.doi.org/10.1037/abn0000258

Layard, R. (2005). *Happiness: Lessons form a new science*. London, UK: Allen Lane.

Marsh, H. W. (1988). Multitrait multimethod analysis. In J. P. Keeves (Ed.), *Educational research methodology, measurement and evaluation: An international handbook* (pp. 570–580). Oxford, England: Pergamon Press.

Marsh, H. W. (1993). Multitrait-multimethod analyses: Inferring each trait-method combination with multiple indicators. *Applied Measurement in Education, 6,* 49–81. http://dx.doi.org/10.1207/s15324818ame0601_4

Marsh, H. W. (2007). *Self-concept theory, measurement and research into practice: The role of self-concept in educational psychology*. Leicester, England: British Psychological Society.

Marsh, H. W., Balla, J., & McDonald, R. P. (1988). Goodness of fit in confirmatory factor analysis: The effect of sample size. *Psychological Bulletin, 103,* 391–410. http://dx.doi.org/10.1037/0033-2909.103.3.391

Marsh, H. W., Ellis, L. A., Parada, R. H., Richards, G., & Heubeck, B. G. (2005). A short version of the Self Description Questionnaire II: Operationalizing criteria for short-form evaluation with new applications of confirmatory factor analyses. *Psychological Assessment, 17,* 81–102. http://dx.doi.org/10.1037/1040-3590.17.1.81

Marsh, H. W., & Grayson, D. (1995). Latent-variable models of multitrait-multimethod data. In R. H. Hoyle (Ed.), *Structural equation modeling: Issues and applications* (pp. 177–198). Thousand Oaks, CA: Sage.

Marsh, H. W., Guo, J., Dicke, T., Parker, P. D., & Craven, R. G. (2019). Confirmatory factor analysis (CFA), exploratory structural equation modeling (ESEM), and Set-ESEM: Optimal balance between goodness of fit and parsimony. *Multivariate Behavioral Research*. Advance online publication. http://dx.doi.org/10.1080/00273171.2019.1602503

Marsh, H. W., Hau, K. T., Artelt, C., Baumert, J., & Peschar, J. L. (2006). OECD's brief self-report measure of educational psychology's most useful affective constructs: Cross-cultural, psychometric comparisons across 25 countries. *International Journal of Testing, 6,* 311–360. http://dx.doi.org/10.1207/s15327574ijt0604_1

Marsh, H. W., Hau, K. T., & Grayson, D. (2005). Goodness of fit evaluation in structural equation modeling. In A. Maydeu-Olivares & J. McArdle (Eds.), *Contemporary psychometrics. A Festschrift for Roderick P. McDonald* (pp. 275–340). Mahwah, NJ: Erlbaum.

Marsh, H. W., Hau, K. T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis testing approaches to setting cutoff values for fit indexes and dangers in overgeneralising Hu & Bentler's (1999) findings. *Structural Equation Modeling, 11,* 320–341. http://dx.doi.org/10.1207/s15328007sem1103_2

Marsh, H. W., & Hocevar, D. (1988). A new, more powerful approach to multitrait-multimethod analyses: Application of second-order confirmatory factor analysis. *Journal of Applied Psychology, 73,* 107–117. http://dx.doi.org/10.1037/0021-9010.73.1.107

Marsh, H. W., Lüdtke, O., Muthén, B., Asparouhov, T., Morin, A. J., Trautwein, U., & Nagengast, B. (2010). A new look at the big five factor structure through exploratory structural equation modeling. *Psychological Assessment, 22,* 471–491. http://dx.doi.org/10.1037/a0019227

Marsh, H. W., Lüdtke, O., Nagengast, B., Morin, A. J., & Von Davier, M. (2013). Why item parcels are (almost) never appropriate: Two wrongs do not make a right—Camouflaging misspecification with item parcels in CFA models. *Psychological Methods, 18,* 257–284. http://dx.doi.org/10.1037/a0032773

Marsh, H. W., Martin, A. J., & Jackson, S. (2010). Introducing a short version of the physical self description questionnaire: New strategies, short-form evaluative criteria, and applications of factor analyses. *Journal of Sport & Exercise Psychology, 32,* 438–482. http://dx.doi.org/10.1123/jsep.32.4.438

Marsh, H. W., Martin, A. J., Yeung, A., & Craven, R. (2017). Competence self-perceptions. In A. J. Elliot, C. Dweck, & D. Yeager (Eds.), *Handbook of competence and motivation* (pp. 85–115). New York, NY: Guilford Press.

Marsh, H. W., Morin, A. J., Parker, P. D., & Kaur, G. (2014). Exploratory structural equation modeling: An integration of the best features of exploratory and confirmatory factor analysis. *Annual Review of Clinical Psychology, 10,* 85–110. http://dx.doi.org/10.1146/annurev-clinpsy-032813-153700

Marsh, H. W., Muthén, B., Asparouhov, T., Lüdtke, O., Robitzsch, A., Morin, A. J., & Trautwein, U. (2009). Exploratory structural equation modeling, integrating CFA and EFA: Application to students' evaluations of university teaching. *Structural Equation Modeling, 16,* 439–476. http://dx.doi.org/10.1080/10705510903008220

Marsh, H. W., Nagengast, B., & Morin, A. J. S. (2013). Measurement invariance of big-five factors over the life span: ESEM tests of gender, age, plasticity, maturity, and la dolce vita effects. *Developmental Psychology, 49,* 1194–1218. http://dx.doi.org/10.1037/a0026913

Marsh, H. W., Pekrun, R., Parker, P., Murayama, K., Guo, J., & Dicke, T. (2019). The Murky Distinction Between Self-Concept and Self-Efficacy Beware of Lurking Jingle-Jangle Fallacies. *Journal of Educational Psychology, 111,* 331–353. http://dx.doi.org/10.1037/edu0000281

Marx, B. P., & Sloan, D. M. (2005). Peritraumatic dissociation and experiential avoidance as predictors of posttraumatic stress symptomatology. *Behaviour Research and Therapy, 43,* 569–583. http://dx.doi.org/10.1016/j.brat.2004.04.004

Meyer, B., Enström, M., Harstveit, M., Bowles, D., & Beevers, C. (2007). Happiness and despair on the catwalk: Need satisfaction, well-being, and personality adjustment among fashion models. *The Journal of Positive Psychology, 2,* 2–17. http://dx.doi.org/10.1080/17439760601076635

Morin, A. J., Marsh, H. W., & Nagengast, B. (2013). Exploratory structural equation modeling. In G. R. Hancock & R. O. Mueller (Eds.), *Structural equation modeling: A second course* (pp. 395–436). Charlotte, NC: Information Age Publishing Inc.

Muthén, L. K., & Muthén, B. O. (2017). *Mplus user's guide* (8th ed.). Los Angeles, CA: Author.

Noetel, M., Ciarrochi, J., Sahdra, B., & Lonsdale, C. (2019). Using genetic algorithms to abbreviate the mindfulness inventory for sport: A

substantive-methodological synthesis. *Psychology of Sport and Exercise, 45,* 101545. http://dx.doi.org/10.1016/j.psychsport.2019.101545

Organisation for Economic Co-Operation and Development. (2013). *OECD guidelines on measuring subjective well-being*. Paris, France: OECD Publishing.

Pejtersen, J. H., Kristensen, T. S., Borg, V., & Bjorner, J. B. (2010). The second version of the Copenhagen Psychosocial Questionnaire. *Scandinavian Journal of Public Health, 38,* 8–24. http://dx.doi.org/10.1177/1403494809349858

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

R Core Team. (2018). *A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/

Roemer, L., Salters, K., Raffa, S. D., & Orsillo, S. M. (2005). Fear and avoidance of internal experiences in GAD: Preliminary tests of a conceptual model. *Cognitive Therapy and Research, 29,* 71–88. http://dx.doi.org/10.1007/s10608-005-1650-2

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

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

Ryan, R. M., & Deci, E. L. (2017). *Self-determination theory: Basic psychological needs in Motivation, development, and wellness*. New York, NY: Guilford Press.

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,* 1069–1081. http://dx.doi.org/10.1037/0022-3514.57.6.1069

Ryff, C. D., & Singer, B. (1996). Psychological well-being: Meaning, measurement, and implications for psychotherapy research. *Psychotherapy and Psychosomatics, 65,* 14–23. http://dx.doi.org/10.1159/000289026

Sahdra, B. K., Ciarrochi, J., Parker, P. D., Marshall, S., & Heaven, P. (2015). Empathy and nonattachment independently predict peer nominations of prosocial behavior of adolescents. *Frontiers in Psychology, 6,* 263. http://dx.doi.org/10.3389/fpsyg.2015.00263

Sahdra, B. K., Ciarrochi, J., Parker, P., & Scrucca, L. (2016). Using genetic algorithms in a large nationally representative American sample to abbreviate the Multidimensional Experiential Avoidance Questionnaire. *Frontiers in Psychology, 7,* 189. http://dx.doi.org/10.3389/fpsyg.2016.00189

Sass, D. A., & Schmitt, T. A. (2010). A comparative investigation of rotation criteria within exploratory factor analysis. *Multivariate Behavioral Research, 45,* 73–103. http://dx.doi.org/10.1080/00273170903504810

Schroeders, U., Wilhelm, O., & Olaru, G. (2016). The influence of item sampling on sex differences in knowledge tests. *Intelligence, 58,* 22–32. http://dx.doi.org/10.1016/j.intell.2016.06.003

Scrucca, L., & Sahdra, B. K. (2015). GAabbreviate: Abbreviating questionnaires (or other measures) using genetic algorithms (Version 1.0: R package) [Computer software]. Retrieved from http://CRAN.Rproject.org/package=GAabbreviate

Self, A., Thomas, J., & Randall, C. (2012). *Measuring national well-being: Life in the UK, 2012*. Retrieved from http://webarchive.nationalarchives.gov.uk/20160105183326/http://www.ons.gov.uk/ons/rel/wellbeing/measuring-national-well-being/first-annual-report-on-measuring-national-well-being/art-measuring-national-well-being-annual-report.html

Seligman, M. E. P. (2002). *Authentic happiness: Using the new positive psychology to realize your potential for lasting fulfillment*. New York, NY: Free Press.

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

Steffick, D. E. (2000). *Documentation of affective functioning measures in the health and retirement study* (Technical Report No. HRS/AHEAD). Ann Arbor, MI: Survey Research Center, University of Michigan.

Stewart-Brown, S., Tennant, A., Tennant, R., Platt, S., Parkinson, J., & Weich, S. (2009). Internal construct validity of the Warwick-Edinburgh Mental Well-being Scale (WEMWBS): A Rasch analysis using data from the Scottish Health Education Population Survey. *Health and Quality of Life Outcomes, 7,* 15. http://dx.doi.org/10.1186/1477-7525-7-15

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,* 251–279. http://dx.doi.org/10.1111/aphw.12027

Tennant, R., Hiller, L., Fishwick, R., Platt, S., Joseph, S., Weich, S., . . . Stewart-Brown, S. (2007). The Warwick-Edinburgh mental well-being scale (WEMWBS): Development and U. K. validation. *Health and Quality of Life Outcomes, 5,* 63. http://dx.doi.org/10.1186/1477-7525-5-63

Thøgersen-Ntoumani, C., & Ntoumanis, N. (2007). A self-determination theory approach to the study of body image concerns, self-presentation and self-perceptions in a sample of aerobic instructors. *Journal of Health Psychology, 12,* 301–315. http://dx.doi.org/10.1177/1359105307074267

Tull, M. T., Gratz, K. L., Salters, K., & Roemer, L. (2004). The role of experiential avoidance in posttraumatic stress symptoms and symptoms of depression, anxiety, and somatization. *Journal of Nervous and Mental Disease, 192,* 754–761. http://dx.doi.org/10.1097/01.nmd.0000144694.30121.89

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,* 1063–1070. http://dx.doi.org/10.1037/0022-3514.54.6.1063

Wei, M., Shaffer, P., Young, S., & Zakalik, R. (2005). Adult attachment, shame, depression, and loneliness: The mediation role of basic psychological needs satisfaction. *Journal of Counseling Psychology, 52,* 591–601. http://dx.doi.org/10.1037/0022-0167.52.4.591

Weinstein, N., & Ryan, R. M. (2010). When helping helps: Autonomous motivation for prosocial behavior and its influence on well-being for the helper and recipient. *Journal of Personality and Social Psychology, 98,* 222–244. http://dx.doi.org/10.1037/a0016984

WHOQoL Group. (1994). The development of the World Health Organization quality of life assessment instrument (the WHOQOL). In J. Orley & W. Kuyken (Eds.), *Quality of life assessment: International perspectives* (pp. 41–57). Berlin, Germany: Springer. http://dx.doi.org/10.1007/978-3-642-79123-9_4

World Health Organization. (1990). *The ICD-10 classification of mental and behavioural disorders: Clinical descriptions and diagnostic guidelines*. Geneva, Switzerland: World Health Organization.

World Health Organization. (2018). *The ICD-11 International statistical classification of diseases and related health problems*. Geneva, Switzerland: World Health Organization.

World Health Organization. (2006). *Official records of the World Health Organization* (Vol. 2). Geneva, Switzerland: World Health Organization. Retrieved from https://apps.who.int/iris/bitstream/handle/10665/85573/Official_record2_eng.pdf?sequence=1

Yarkoni, T. (2010). The abbreviation of personality, or how to measure 200 personality scales with 200 items. *Journal of Research in Personality, 44,* 180–198. http://dx.doi.org/10.1016/j.jrp.2010.01.002

Received April 15, 2019
Revision received September 3, 2019
Accepted October 18, 2019 ■