Current recommendations on the selection of measures for well-being

Tyler J. VanderWeele
Harvard T.H. Chan School of Public Health,
Boston, MA, U.S.A.
tvanderw@hsph.harvard.edu

Claudia Trudel-Fitzgerald
Harvard T.H. Chan School of Public Health
Boston, MA, U.S.A.
ctrudel@hsph.harvard.edu

Paul Allin
Imperial College London
London, U.K.
p.allin@imperial.ac.uk

Colin Farrelly
Queen's University
Kingston, ON, Canada
farrelly@queensu.ca

Guy Fletcher
University of Edinburgh
Edinburgh, U.K.
Guy.Fletcher@ed.ac.uk

Donald E. Frederick
Harvard University
Cambridge, MA, U.S.A.
donald.frederick@gmail.com

Jon Hall
United Nations Development Programme
New York, NY, U.S.A.
jonathan.hall@undp.org

John F. Helliwell

University of British Columbia
Vancouver, BC, Canada
johnfhelliwell@gmail.com

Eric S. Kim
Harvard T.H. Chan School of Public Health
Boston, MA, U.S.A.

eskim@hsph.harvard.edu

William A. Lauinger
Chestnut Hill College,
Philadlphia, PA, U.S.A.
LauingerW@chc.edu

Matthew T. Lee
Harvard University
Cambridge, MA, U.S.A.
matthew_lee@fas.harvard.edu

Sonja Lyubomirsky
University of California, Riverside
Riverside, CA, U.S.A.
sonja.lyubomirsky@ucr.edu

Seth Margolis
University of California, Riverside
Riverside, CA, U.S.A.
sethmmargolis@gmail.com

Eileen McNeely
Harvard T.H. Chan School of Public Health
Boston, MA, U.S.A.
emcneely@hsph.harvard.edu

Neil Messer
University of Winchester
Winchester, U.K.
Neil.Messer@winchester.ac.uk

Louis Tay
Purdue University
West Lafayette, IN, U.S.A.
stay@purdue.edu

Vish Viswanath

Harvard T.H. Chan School of Public Health
Boston, MA, U.S.A.
vish_viswanath@dfci.harvard.edu

Dorota Węziak-Białowolska
Harvard T.H. Chan School of Public Health
Boston, MA, U.S.A.
doweziak@hsph.harvard.edu

Laura D. Kubzansky
Harvard T.H. Chan School of Public Health
Boston, MA, U.S.A.
lkubzans@hsph.harvard.edu

Abstract

Measures of well-being have proliferated over the past decades. Very little guidance has been available as to which measures to use in what contexts. This paper provides a series of recommendations, based on the present state of knowledge and the existing measures available, of what measures might be preferred in which contexts. The recommendations came out of an interdisciplinary workshop on the measurement of well-being. The recommendations are shaped around the number of items that can be included in a survey, and also based on the differing potential contexts and purposes of data collection such as, for example, government surveys, or multi-use cohort studies, or studies specifically about psychological well-being. The recommendations are not intended to be definitive, but to stimulate discussion and refinement, and to provide guidance to those relatively new to the study of well-being.

Word Count for Abstract: 137
Word Count for Text (excluding in text citations): 4,443

Introduction

Over the last several years, interest in the measurement and promotion of well-being has increased exponentially with calls for societal transformation and a new vision for health that places well-being at the center (Plough, 2015). As research on well-being both as an outcome (or a target for monitoring) and as a predictor of other health-related outcomes has expanded dramatically, conceptions and measures of well-being have likewise proliferated. Consequently, it can be challenging to compare ideas and findings across different measures and conceptions. For example, well-being can be characterized by objective measures, also referred to as measures related to "standard of living," and by subjective measures, based on cognitive and affective judgements a person makes about their life (Stiglitz et al., 2010). Objective aspects of well-being will of course also influence subjective well-being levels (Patel et al., 2018). Many countries routinely collect data on various factors that are considered indicators of objective well-being, including measures of educational attainment, safety, income, life expectancy, and so forth. Only a few countries have begun to collect data on subjective well-being measures, notably life satisfaction and happiness, on a regular basis. However, subjective well-being has been an important area of research in psychology for decades (Myers and Diener, 2018), and increasingly in other academic disciplines as well (Ngamaba et al., 2018). Subjective well-being is moreover not merely the absence of mental illness; indeed measures of subjective well-being predict strongly and independently subsequent mental illness above and beyond baseline measures of mental illness (Wood and Joseph, 2010). The focus of the present paper is to provide a set of recommendations concerning measuring subjective well-being.

At least three conceptual approaches to evaluating subjective well-being are commonly used, including hedonic, evaluative, and eudaimonic conceptions of well-being (Ryff, 1989; Kahneman et al., 2003; National Research Council, 2013; OECD, 2013). A hedonic perspective focuses on well-being understood as whether one feels happy or experiences pleasure and lacks pain; an evaluative perspective focuses on well-being defined by one's view of, or overall satisfaction with life or different domains of life. Closely related desire-fulfillment theories, while receiving considerable attention in philosophy (Fletcher, 2016) have only recently been empirically operationalized (Margolis et al., 2019). An eudaimonic perspective focuses on whether individuals feel they have attained self-realization, or if they are fully-functioning or

fulfilling a sense of purpose. There is general agreement that well-being itself is a broad multi-dimensional construct that extends beyond simply feeling happy or being satisfied with life (Stiglitz et al., 2017; OECD, 2019).

A distinction might also be drawn between "psychological well-being," which concerns assessment of an individual's various psychological states, versus "subjective well-being," which includes an individual's subjective assessment of any aspect of their life (e.g., finances, physical health). While these two terms are often used interchangeably, neither of these categories encompasses the other. Psychological states can be assessed by direct observation (e.g., of textual communication) rather than subjective self-report; conversely, one can report subjectively on states that are not psychological, like one's physical health. In this paper, we will use "psychological well-being" or "subjective well-being" in their more general descriptive senses rather than to refer to any specific measures.

Despite measurement and conceptual challenges, the recent proliferation of studies on well-being has provided an exciting array of results and novel insights. However, there is, as yet, little guidance as to what to measure or which scale to use for any particular investigation. Answers to these kinds of questions inevitably depend on the context, the resources available, and the goals of measurement. Different measures may be better suited to studying the determinants of well-being, versus understanding the effects of aspects of well-being on other outcomes. Although a number of good overviews of different subjective well-being measures are available (National Research Council, 2013; OECD, 2013; Su et al., 2014; Tay et al., 2014; Hone et al., 2014), these generally provide a compendium of existing measures (or information on where to find the measures) rather than specific guidance regarding which measures to use in what contexts. In this paper, we put forward a series of recommendations for selecting measures of subjective well-being across different contexts, focusing on the use of subjective, rather than objective, measures.

Methods

These recommendations arose out of an interdisciplinary workshop on the measurement of well-being hosted at Harvard University in April 2018, drawing upon a multidisciplinary group of well-being experts from around the world. Discussions of well-being measures are often confined to experts within a single discipline. However, greater understanding of both the science of well-being and the measurement issues may be gained from considering research on

well-being in studies across multiple disciplines. Thus, the workshop conveners, directors of the Lee Kum Sheung Center for Health and Happiness and of the Human Flourishing Program (both at Harvard University), identified workshop participants by seeking scholars who have actively contributed to the study of well-being and who also, together, could broadly represent numerous disciplines including psychology, sociology, economics, political science, public health, medicine, statistics, philosophy, and theology. In addition to individual workshop presentations and discussions about the study of well-being from a range of disciplinary perspectives, several sessions were devoted to questions of measurement recommendations. Building on these discussions, an initial set of recommendations was drafted, and all workshop participants were invited to comment and contribute. Final recommendations, further refined in subsequent discussion and written exchange, are presented here.

A key component driving the discussion and ensuing recommendations was recognition that there will not be a one-size-fits-all recommendation for measuring well-being. Recommendations must be informed by careful consideration of each type of research or reporting endeavor and the likely constraints on the number of items that can be used to measure well-being in specific contexts (e.g., government surveys, multi-use cohort studies; studies specifically about well-being). Thus, different recommendations are made depending on the purpose for which a well-being measure is sought, with rationale provided for the choice of measures. These recommendations are not intended to be definitive, but rather constitute the consensus of our interdisciplinary panel of experts, given the present state of knowledge and the measures currently available. Our goals are to provide practical guidance for the present moment and to stimulate debate and discussion, which we expect will refine well-being measurement further as new research in this area emerges. Recommendations are organized according to their intended use and, within each section, giving consideration to the options available depending on the number of items a given project might be able to accommodate.

Results

Psychological Well-Being in Government Surveys

Government surveys are frequently designed for the purposes of monitoring and surveillance. Our recommendations for assessing psychological well-being in government surveys with a very limited number of items follows that of the United Kingdom's Measuring

National Well-Being Programme. In 2010, the UK government committed to assessing national levels of well-being (Allin and Hand, 2017). To accomplish this, the UK's Office for National Statistics established a Measuring National Well-Being Programme to identify key areas that mattered most to people and to make an initial proposal for domains and specific measures. This Programme drew upon existing frameworks in the well-being literature, including prior work by the OECD (Hall et al., 2010), and aimed to incorporate items for subjective well-being already used in the international well-being literature. They included items related to hedonic, evaluative, and eudaimonic well-being, but also tried to keep the number of questions limited to avoid excessive costs and enable widespread use. The questions were tested in the Annual Population Survey of households, and a final set of four questions has now been included on the annual UK National Survey since 2011 (Allin and Hand, 2017). Based on the thoughtful engagement of the UK's Measuring National Well-Being Programme, the choice of questions already widely used in well-being research, the range of questions administered, and the successful record of data collection on these questions, we recommend using this same four-question set for obtaining a brief assessment of psychological well-being via government surveys or other large-scale population-wide monitoring instrument. Although other countries and organizations have also included additional well-being questions, such as those included in the European Social Survey, these constitute a much longer list of questions and may be less suitable for very brief well-being assessments. The four questions from the UK National Survey are:

1. Overall, how satisfied are you with life as a whole these days?
2. Overall, to what extent do you feel the things you do in your life are worthwhile?
3. Overall, how happy did you feel yesterday?
4. Overall, how anxious did you feel yesterday?

Questions are asked using a 0–10 response scale where 0 is "Not at all" and 10 is "Completely." This limited set are easily incorporated into existing surveys and relatively quick to administer. Moreover, any monitoring body using these items could immediately compare their findings with UK statistics. The four questions draw from each of the broad conceptual approaches to psychological well-being: evaluative well-being (item #1 [life satisfaction]), eudaimonic well-being (item #2 [purpose/meaning in life]), and hedonic well-being and ill-

being, respectively (items #3 [positive affect], #4 [negative affect]). Gallup, OECD, and other large-scale organizations engaged in monitoring subjective well-being also use the above items in assessing evaluative and eudaimonic well-being (i.e., life satisfaction and worthwhile activities, respectively; OECD, 2013). The items evaluating hedonic well-being and ill-being query positive and negative affect, respectively, and sample from the person's experience the prior day. While enquiring only about a single day may not be representative of life more broadly and is perhaps less suitable for etiologic research purposes, it does provide an assessment of positive and negative affect for the country or region as a whole when responses are averaged over numerous persons on different days (Allin and Hand, 2017). Thus, they may be useful for monitoring and tracking. While some have aggregated the four questions by taking a sum score across the items (Benson et al., 2019), items represent distinct conceptual domains and are generally reported separately.

When even four items are too many to include on a given survey, for an even briefer two-item survey, we recommend assessing evaluative and eudaimonic well-being using the life satisfaction (#1) and worthwhile activity (#2) questions. These two items have been used extensively, have broad conceptual coverage, and, across numerous individual items, show some of the highest and most consistent correlations with much broader well-being measures (Cheung and Lucas, 2014; Helliwell et al., 2016; OECD, 2013). When it is possible to include only a single item, we recommend assessing evaluative well-being (item #1). Although measuring life satisfaction alone is subject to numerous limitations (Kahneman et al., 2003; Allin and Hand, 2017; VanderWeele, 2017; Ryff, 1989), if only one question can be included, life satisfaction does provide a relatively broad assessment and has been found to perform similarly compared to multiple-item life satisfaction scales in prior work (Cheung and Lucas, 2014). Moreover, this item has been used in surveys around the world (Helliwell et al., 2016), which allows comparisons across countries. For more substantial assessments, perhaps also targeted not only for monitoring but also for research, see also the sections below.

Psychological Well-Being in Multi-Use Cohort Studies

Increasingly, multi-purpose cohort studies have been seeking well-being items to include in data collection instruments for use with explanatory research (rather than monitoring and surveillance) that may examine well-being either as an outcome (i.e., dependent variable) or as a predictor (i.e., independent variable/exposure) of other outcomes. When considering well-being

as an outcome, a broader conceptualization can be appropriate, but specific aspects of well-being can also be examined. When considering well-being as a predictor of other health-related outcomes, more specific conceptualizations are likely to be useful, with a particular focus on items that predict future changes in health and behavior. For many multi-use cohort studies, space constraints often make it possible to include only a handful of items. In these circumstances, we recommend the following six questions drawn from the evaluative, eudaimonic, hedonic, and other domains. The items could be used as predictors or as outcomes in etiologic research.

1. Overall, how satisfied are you with life as a whole these days?
2. Overall, to what extent do you feel the things you do in your life are worthwhile?
3. In general, how happy or unhappy do you usually feel?
4. I have a sense of direction and purpose in life
5. Overall, I expect more good things to happen to me than bad
6. If something can go wrong for me, it will (reverse coded)

The first two items are scored from 0 = "Not at all" to 10 = "Completely". The third item is scored from 0 ("Extremely Unhappy") to 10 ("Extremely Happy)" and the fourth item from 0 ("Strongly Disagree") to 10 ("Strongly Agree"). The fifth and sixth items have traditionally been scored from 1 ("Strongly Disagree") to 5 ("Strongly Agree"), but could also be scored from 0-10 for consistency with the others.

Including the first two questions has the advantages discussed above. For item #3, unlike the question used in the UK survey, which asks about happiness level on the previous day, the question here is phrased according to general levels of happiness (Fordyce, 1988). This may be more suitable for individual level etiologic research purposes because it captures a more stable, enduring experience in contexts where well-being is inquired only sporadically (Hudson et al., 2017). Among the various dimensions of psychological well-being, purpose and optimism are among those that are most consistently and strongly related to physical health outcomes, including all-cause mortality in prospective studies (Trudel-Fitzgerald et al., 2019); thus, we suggest two questions to capture each of these domains. The optimism items (#5 and #6) are drawn from the Life Orientation Test-Revised (LOT-R; Scheier et al., 1994) using the items most

predictive of mortality. For purpose, the worthwhile activities item (#2) is supplemented by an item (#4) from the purpose subscale of the Psychological Well-being Scale (Ryff, 1989).

Studies of Psychological Well-Being

The greatest progress in the science of well-being will likely come from large studies designed specifically to measure and study well-being itself. For this purpose, we recommend scales and inventories that include assessment of multiple aspects of psychological well-being, including life satisfaction, positive affect, meaning, purpose, and personal growth, among others. Some have argued that composite measures of well-being that aggregate across these various dimensions can be useful in gaining a broad perspective on potential determinants of overall well-being, and might be valuable as a focus for policy (e.g., Su et al., 2014). Evidence suggests that the overall aggregates of various different multi-dimensional well-being scales are themselves often strongly correlated (Goodman et al., 2018) and thus contain very similar information, though, if dichotomized, differing dichotomization schemes can of course lead to different conclusions concerning e.g. prevalence (Hone et al., 2014). From a scientific perspective, however, when seeking to understand the causes and consequences of distinct aspects of psychological well-being, the use of more specific measures is necessary. In fact, different dimensions of psychological well-being very likely have different causes and different effects (Baumeister et al., 2013; Trudel-Fitzgerald et al., 2019). It is thus the specific dimensions included within a scale or inventory that will likely be most relevant for the scientific study of well-being, since aggregate measures are often be similar. This perspective shapes the remainder of the recommendations in this paper. Various validated scales that measure specific dimensions of psychological well-being, such as those developed by Diener et al. (1985), Ryff (1989), Lyubomirsky and Lepper (1999), Keyes (2002), Su et al. (2014), Warwick Medical School (2018), and others, might be used for this purpose.

When seeking to study specific dimensions of psychological well-being, we recommend, when possible, the use of at least two different scales designed to assess the same construct as a sensitivity analysis for the robustness of the conclusions being drawn. For example, for meaning and purpose, one might use both the purpose subscale from the Psychological Well-being Scale (Ryff, 1989) and Meaning in Life questionnaire (Steger et al., 2006) in the same study. Such practice may also help address aspects of measurement that have not yet, or only recently, been

adequately conceptualized. For instance, while meaning and purpose in life are often combined in measures designed to capture a single construct, recent empirical and conceptual work has suggested three distinct facets (Martela and Steger, 2016). The use of more than one scale to assess the same construct may help facilitate such insights.

If a study seeks to examine numerous domains of psychological well-being, either as a predictor or as an outcome, then a broad multi-dimensional inventory will most likely be desirable, because such a measure can be considered either as a single composite or by specific sub-domains. Among the existing available measures that include multiple items for many different dimensions, we recommend the 54-item Comprehensive Inventory of Thriving (CIT; Su et al., 2014). This inventory was created based on a prior survey of other multi-dimensional approaches to and measures of psychological well-being (e.g., Diener et al., 1985, 2009; Ryff, 1989; Scheier et al., 1994; Seligman, 2011), and includes multiple items per dimension. The CIT includes three items each for 18 facets that are grouped within the following seven dimensions: relationships (support, community, trust, respect, loneliness, belonging), engagement, mastery (skills, learning, accomplishment, self-efficacy, self-worth), autonomy, meaning, optimism, and subjective well-being (life satisfaction, positive feelings, absence of negative feelings) (α= 0.71 to 0.96 across varied populations; Su et al., 2014). Its psychometric properties and measurement invariance have also been examined in cross-cultural settings (Wiese et al., 2018). Once again, we recommend that, if possible, the study of each CIT construct be supplemented by the use of other scales (e.g., Diener et al., 1985; Ryff, 1989; Scheier et al., 1994; Lyubomirsky and Lepper, 1999; Keyes, 2002; Steger et al., 2006; Martela and Steger, 2016) purportedly assessing the same construct.

Human Flourishing

Human flourishing or complete human well-being is the broadest possible construct under the study of well-being. Notably, it has been conceptualized as "the achievement of all goods, purposes and ends of human existence" (Messer, 2013) or as "a state in which all aspects of a person's life are good" (VanderWeele, 2017). Such ends and goods include not only psychological well-being but also physical health, a domain that is absent from many of the scales discussed above, and character, and could also include both objective and subjective assessments. As before, we will focus here on the subjective aspects. Important to note, is that

because it is so broad, the construct of flourishing should ideally capture, amongst other things, multiple facets of psychological well-being (e.g., hedonic, evaluative, and eudaimonic), as relevant sub-components of the larger experience. As such, in many research contexts, "flourishing" makes sense principally as an outcome, rather than as a predictor. It would make little sense to examine the effects of flourishing on subsequent physical health if the flourishing construct itself includes physical health. However, assessing flourishing is useful in other contexts (e.g., examining the effect of individual employee flourishing on various objective outcomes including productivity or turnover).

Developing valid measures is a complex process, especially when the construct is as broad as flourishing. There may be tension between capturing as many domains as possible versus the danger of including domains that are relatively less important or trivial. A focus on those dimensions of human well-being that are ends in themselves and nearly universally desired may help shape consensus on what to measure (VanderWeele, 2017). A number of conceptualizations and measures of flourishing have been developed (Keyes, 2002; Diener et al., 2009; Seligman, 2011; Huppert and So, 2013; Hone, 2014; VanderWeele, 2017), though many of these do not include physical health. To enhance reliability and to make it possible to consider various dimensions separately, we believe at least two or three items per domain assessed would be desirable. Several existing approaches make use of one only item per domain (Diener et al., 2009; Huppert and So, 2013). For a longer multi-item comprehensive assessment of subjective flourishing, because of its breadth, as noted above, we recommend supplementing Su et al.'s (2014) Comprehensive Inventory of Thriving (CIT) described above, which covers multiple dimensions of psychological well-being, with a multi-item assessment of physical health such as the 12-Item Short-Form Health Survey (SF-12; Ware et al., 1996). The SF-12 is widely used, has demonstrated good psychometric properties (e.g., 2-week test-retest reliability, r=0.86 in UK adults and 0.89 in US adults, [Ware et al.,1996]; α= 0.70 to 0.89 across samples of older adults, [Resnick and Nahm, 2001]), and captures the dimension of physical health that is absent from the CIT. For a brief 10-item flourishing measure that may also permit separate consideration of domains, we recommend VanderWeele's (2017) Flourishing Index, which comprises two items for each of the following domains: happiness and life satisfaction, mental and physical health, meaning and purpose, character and virtue, and close social relationships. These items were chosen from among those most commonly used and previously validated prior well-being scales.

The scale has had some degree of empirical validation (α=0.89, Węziak-Białowolska et al., 2019a) and its psychometric properties and measurement invariance have also been recently examined in cross-cultural settings (Węziak-Białowolska et al., 2019b).

Discussion

As noted above, our recommendations are provisional, drawing on the current state of knowledge and the existing validated measures available. Although recommending a set of validated items, as in many cases above, is not comparable to the process of validating a set of items when combined into a new measure, we hope these recommendations will help facilitate subsequent research on well-being and on its measurement. Below, we consider a number of other future developments that may further improve our ability to measure, study, and track well-being.

In the discussion above, for settings in which only a single well-being item will be used, we recommended the question, "Overall, how satisfied are you with life as a whole these days?" While investigators have generally referred to this item as a "cognitive" or "evaluative" measure of psychological well-being, it does place strong emphasis on satisfaction, rather than on whether all facets of life are in fact good. This could be problematic. A person can be satisfied and addicted to narcotics, or satisfied and completely socially isolated. It is not clear that it is reasonable in such cases to say that human well-being is high. Although unusual examples, they demonstrate the potential that assessing life satisfaction alone which, out of context, may not represent an accurate portrait of well-being or flourishing in life as a whole. This critique may be less relevant to other forms of life evaluation that do not make explicit reference to "satisfaction," including the Cantril Ladder (Helliwell et al., 2016). However, the latter requires considerably more space than the simple life satisfaction question, but may be preferable if there are not strict constraints on space. Other single item measures that might more holistically consider self-report evaluations of well-being across the whole of life, and that are less focused only on the satisfaction of desires might deserve further study, such as "All aspects of my life at present are good," or "All is well with my life." Whether they perform better than the widely-studied life satisfaction item above requires further research and assessment. Until better studied, it may be desirable, when possible, to include at least two single-item overall evaluative measures.

While existing measures capture a number of important dimensions, other aspects of well-being are absent, as pointed by philosophers and others (e.g., Fletcher, 2016). First, few well-being scales make any attempt to capture the value of existing knowledge or processes necessary for acquiring it. Second, most scales focus almost exclusively on individual well-being. Although some measures include items assessing the quality of an individual's social relationships, broader community well-being is often overlooked. Examining community well-being (e.g., within a family, city, or nation) may also be important for a broader understanding of the determinants and consequences of individual well-being (Allin and Hand, 2017; Phillips and Wong, 2017). It may thus be useful to supplement measures of individual well-being, and their aggregates, with measures of community well-being (Phillips and Wong, 2017; VanderWeele, 2019; Allin and Hand, 2017).

Third, although some measures of *spiritual* well-being are available (Paloutzian and Ellison, 1982; Peterman et al., 2002), the most widely used *general* well-being scales do not capture spiritual well-being. This is potentially problematic. For much of the world's population, some notion of spirituality or religion is highly important (Pew Religious Landscape Study, 2018; Diener et al., 2011), and many consider it the most important aspect of well-being. Including spiritual well-being items within general well-being scales is challenging because, to a greater extent than with other aspects of well-being, the way in which this construct is understood likely varies across religious and spiritual traditions. Thus, tradition-specific measures of spiritual well-being may be an important and necessary step forward (VanderWeele et al., 2019). Such measures could potentially supplement more generic and universal well-being measures.

Fourth, although many psychological well-being scales include some notion of autonomy, they are often framed negatively and principally assess whether individuals feel they can make decisions free from influence of others. While useful in many contexts, in some cultures, this formulation may be considered relatively less essential to well-being. The existing measures, moreover, often do not capture positive notions of having freedom to pursue what is important in life. Existing negatively-framed autonomy scales might thus be supplemented with an item like "I am free to pursue what is most important" or like Gallup's question, "Are you satisfied or dissatisfied with your freedom to choose what you do with your life?" Such items may also help to capture aspects of well-being that are important to some individuals but not to

others. For example, artists who cannot pursue artistic expression, creation, and aesthetic experience, may feel their well-being is severely compromised. However, for others, the absence of art for a time, even if they enjoy it, may not similarly substantially compromise well-being. An item such as "I am free to pursue what is most important" may help address these nuances. Such ideas arguably bear some correspondence to Sen's capabilities approach to well-being (Sen, 1999), although its empirical operationalizations have tended to focus on more objective measures (Alkire, 2002; Alkire and Santos, 2010).

Fifth, more work could be done examining important cross-cultural variations in what aspects of well-being are considered most important in different contexts, and whether current measures of well-being, mostly developed in western and high-income countries, may be missing other elements important in other cultures. Future well-being measure development and refinement might consider these potential omissions.

Conclusions

The recommendations in this paper are not intended to be definitive but rather to i) provide guidance for those needing to make practical decisions about well-being measurement today, and ii) prompt further discussion and debate that will eventually lead to further refinement. That well-being is measured—and how it is measured—is critical. What investigators, practitioners, and policy makers measure shapes what they discuss, what priorities they set, and what they aim for. Studies to advance our understanding of the distribution, determinants, and consequences of well-being are essential in efforts to try to improve well-being. However, such studies cannot take place without proper measurement, which in turn is shaped by the purposes and constraints (e.g., regarding number of items) of any particular study. If the well-being of individuals and nations does not get measured, then the focus will likely shift to other indicators, such as only income or physical health. We hope that the recommendations offered here might facilitate more frequent, effective, and impactful measurement of well-being.

Author Contributions

This set of recommendations came about as a result of an interdisciplinary workshop on the measurement of well-being of which the authors were a part. The critical intellectual content of the paper was shaped by all authors and would not have been possible without their participation. The paper represents the collective thought of the author list. The paper was drafted by Tyler J. VanderWeele, Claudia Trudel-Fitzgerald, and Laura D. Kubzansky. All authors provided critical review, feedback, additions, further references, and guidance.

Acknowledgements

The research was supported by the John Templeton Foundation [grant 61075] to Dr. Tyler J. VanderWeele and by funding from the Lee Kum Sheung Center for Health and Happiness at the Harvard T. H. Chan School of Public Health and the Human Flourishing Program at Harvard University..

Declaration of interests

Tyler J. VanderWeele and Eileen McNeely report having previously received consulting fees from Aetna Inc.

References

Alkire S. Valuing Freedoms: Sen's Capability Approach and Poverty Reduction. Oxford University Press,:Oxford, UK, 2002.

Alkire S, Santos ME. Acute multidimensional poverty: A new index for developing countries. United Nations Development Programme Human Development Report Office Background Paper. 2010 Jul 1(2010/11).

Allin P, Hand DJ. New statistics for old? - measuring the wellbeing of the UK. Journal of the Royal Statistical Society, Series A, 2017; 180:1-22.

Baumeister RF, Vohs KD, Aaker JL, Garbinsky EN. Some key differences between a happy life and a meaningful life. The Journal of Positive Psychology, 2013; 8(6), 505-516.

Benson T, Sladen J, Liles A, Potts HWW. Personal Wellbeing Score (PWS)—a short version of ONS4: development and validation in social prescribing. BMJ Open Quality 2019;8:e000394. doi: 10.1136/bmjoq-2018-000394

Cheung F, Lucas RE. Assessing the validity of single-item life satisfaction measures: Results from three large samples. Qual Life Res 2014; 23(10): 2809-2818.

Diener, E., Wirtz, D., Tov, W., Kim-Prieto, C., Choi, D., Oishi, S. et al. (2009). New measures of well-being: Flourishing and positive and negative feelings. Social Indicators Research, 39, 247–266.

Diener E, Emmons RA, Larsen RJ, Griffin S. The satisfaction with life scale. Journal of Personality Assessment, 1985; 49:71–75.

Fletcher G. The Routledge Handbook of Philosophy of Well-Being. Routledge. 2016.

Fordyce MW. A review of research on the happiness measures: A sixty second index of happiness and mental health. Soc Indic Res 1988; 20:355–381.

Goodman FR, Disabato DJ, Kashdan TB, Kauffman SB. Measuring well-being: A comparison of subjective well-being and PERMA. The Journal of Positive Psychology. 2018 Jul 4;13(4):321-32.

Hall J, Giovannini E, Morrone A, Ranuzzi G. A framework to measure the progress of societies. Working Paper 34, STD/DOC(2010)5. Statistics Directorate, Organisation for Economic Co-operation and Development, Paris, 2010 (Available at: http://kniknowledgebase.org/wp-content/uploads/2015/08/A_Framework_to_Measure_the_Progress_of_Societies.pdf) Last Accessed: June 12, 2019

Helliwell J, Layard R, Sachs, J. World Happiness Report 2016, Update (Vol. I). New York: Sustainable Development Solutions Network, 2016.

Hone LC, Jarden A, Schofield GM, Duncan S. Measuring flourishing: The impact of operational definitions on the prevalence of high levels of wellbeing. International Journal of Wellbeing, 2014; 4(1), 62-90.

Hudson NW et al. Day-to-day affect is surprisingly stable: A two-year longitudinal study of well-being. Soc Psychol Personal Sci 2017; 8(1): 45-54.

Huppert FA, So TT. Flourishing across Europe: Application of a new conceptual framework for defining well-being. Social indicators research. 2013 Feb 1;110(3):837-61.

Kahneman D, Diener E, Schwarz N. Well-being: the Foundations of Hedonic Psychology. Russell Sage Foundation: New York, NY, 2003.

Keyes CL. The mental health continuum: from languishing to flourishing in life. J Health Soc Behav, 2002; 43(2), 207-222.

Lyubomirsky S, Lepper H. A measure of subjective happiness: Preliminary reliability and construct validation. Social Indicators Research, 1999; 46:137-155.

Margolis S, Schwitzgebel E, Ozer DJ, Lyubomirsky S. Empirical relationships among five types of well-being. In: M. Lee, L.D. Kubzansky, and T.J. VanderWeele. Measuring Well-Being: Interdisciplinary Perspectives from the Social Sciences and the Humanities. Oxford University Press, forthcoming, 2019.

Martela F, Steger MF. The three meanings of meaning in life: Distinguishing coherence, purpose, and significance. The Journal of Positive Psychology, 2016; 11:5, 531-545.

McKnight PE, Kashdan TB. Purpose in life as a system that creates and sustains health and well-being: An integrative, testable theory. Review of General Psychology, 2009; 13(3), 242-251.

Messer N. Flourishing: Health, disease and bioethics in theological perspective. Eerdmans Publishing: Grand Rapids, MI, 2013.

Myers DG, Diener E. The scientific pursuit of happiness. Perspectives on Psychological Science, 2018;13(2): 218-225.

National Research Council. Subjective Well-Being (National Academies Press, Washington, DC), 2013.

Ngamaba KH. Income inequality and subjective well-being: A systematic review and meta-analysis." Qual Life Res 2018; 27(3): 577-596.

Paloutzian RF, Ellison CW. Loneliness, spiritual well-being, and the quality of life. In L.A. Peplau, D. Perlman (eds.). Loneliness: A Sourcebook of Current Theory, Research and Therapy. New York: John Wiley & Sons, pp 224-237, 1982.

Patel V., et al. The Lancet Commission on global mental health and sustainable development. Lancet 392(10157): 2018; 1553-1598.

Peterman AH, Fitchett G, Brady MJ, Hernandez L, Cella D. Measuring spiritual well-being in people with cancer: The Functional Assessment of Chronic Illness Therapy–Spiritual Well-Being Scale (FACIT–Sp) Annals of Behavioral Medicine, 2002; 24:49–58

Phillips R, Wong C. Handbook of Community Well-Being Research. Springer: Dordrecht, 2017.

Plough AL. Building a culture of health: a critical role for public health services and systems research. Am J Public Health. 2015; 105 Suppl 2:S150-2.

Stiglitz JE, Sen A, Fitoussi J-P. Mismeasuring Our Lives. New York: The New Press, 2010.

OECD Guidelines on Measuring Subjective Well-Being (OECD Publishing, Paris), 2013.

OECD. How's life?. Paris: OECD Publishing, 2019.

Resnick B, Nahm ES. Reliability and validity testing of the revised 12-item Short-Form Health Survey in older adults. J Nurs Meas 2001; 9(2): 151-161.

Ryff CD. Happiness is everything, or is it? Explorations on the meaning of psychological well-being. J Pers Soc Psychol 1989; 57:1069–1081.

Scheier MF. et al. Distinguishing optimism from neuroticism (and trait anxiety, self-mastery, and self-esteem): a reevaluation of the Life Orientation Test. J Pers Soc Psychol 1994;67(6): 1063-1078.

Seligman MEP. Flourish: A visionary new understanding of happiness and well-being. New York: Free Press, 2011.

Sen A. Development as Freedom. New York, Knopf Press, 1999.

Steger MF, Frazier P, Oishi S, Kaler M. The Meaning in Life Questionnaire: Assessing the presence of and search for meaning in life. J Couns Psychol 2006; 53:80–93.

Stiglitz JE, Sen A, Fitoussi JP. Report by the commission on the measurement of economic performance and social progress, 2017.

Su R, Tay L, Diener E. The development and validation of the comprehensive inventory of thriving (CIT) and the brief inventory of thriving (BIT). Appl Psychol Health Well-Being 2014; 6:251–279.

Tay L, Chan D, Diener E. The metrics of societal happiness. Social Indicators Research, 2014;117, 577-600.

Trudel-Fitzgerald C., Millstein R., von Hippel C., Howe R, Tomasso LP, Wagner G, VanderWeele TJ. Psychological well-being as part of the public health debate? Insight into dimensions, interventions, and policy. BMC Public Health, in press.

VanderWeele TJ. On the promotion of human flourishing. Proceedings of the National Academy of Sciences, U.S.A., 2017;31:8148-8156.

VanderWeele TJ, McNeely E, Koh HK. Reimagining health: flourishing. Journal of the American Medical Association, 2019; 321(17):1667-1668.

VanderWeele TJ. Measures of community well-being: a template. International Journal of Community Well-Being, in press. Published Online Early: doi: 10.1007/s42413-019-00036-8

VanderWeele, TJ, Long K, Balboni MJ. Tradition-specific measures of spiritual well-being. In: M. Lee, L.D. Kubzansky, and T.J. VanderWeele. Measuring Well-Being: Interdisciplinary Perspectives from the Social Sciences and the Humanities. Oxford University Press, forthcoming, 2019.

Ware JE, Kosinski M, Keller SD 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. Medical Care, 1996; 34:220-233.

Warwick Medical School. Warwick-Edinburgh Mental Wellbeing Scale (2018 version) at https://warwick.ac.uk/fac/sci/med/research/platform/wemwbs. Last Accessed: June 12, 2019.

Wiese CW, Tay L, Su R, Diener E. Measuring thriving across nations: Examining the measurement equivalence of the Comprehensive Inventory of Thriving (CIT) and the Brief Inventory of Thriving (BIT). Applied Psychology: Health and Well-Being. 2018 Mar;10(1):127-48.

Węziak-Białowolska D, McNeely E, VanderWeele TJ. Flourish index and secure flourish index – validation in workplace settings. Cogent Psychology, 2019a; 6 (1598926): 1-10.

Węziak-Białowolska D, McNeely E, VanderWeele TJ. Human flourishing in cross cultural settings: evidence from the US, China, Sri Lanka, Cambodia and Mexico. Frontiers in Psychology, 2019b; 10 (Article 1269): 1-13.

Wood AM, Joseph S. The absence of positive psychological (eudemonic) well-being as a risk factor for depression: A ten year cohort study. Journal of affective disorders. 2010; 122(3):213-217.