Commentary on “The protective effects of wellbeing and flourishing on long-term mental health risk”

Dorota Weziak-Bialowolska a, b, *

a Sustainability and Health Initiative (SHINE), Department of Environmental Health, Harvard T. H. Chan School of Public Health, 655 Huntington Ave, G28, Boston, MA, USA
b Human Flourishing Program, Institute for Quantitative Social Science, Harvard University, 129 Mount Auburn Street, Cambridge, MA, 02138, USA

ARTICLE INFO

Keywords
Flourishing
Well-being
Instrument
Psychometrics

The concept of human flourishing has been gaining attention over the recent years. Three distinct perspectives have been presented by now: (1) human flourishing as a determinant of health, (2) human flourishing as an outcome of health, and (3) human flourishing as an extension of the definition of health (Levin, 2021). Although discussions on these perspectives continue and the consensus seems distant, there is a growing conviction that human flourishing is a concept that needs to be studied.

The article by Burns et al. (2022) in this issue of ‘Social Science and Medicine – Mental Health’ fits into the first perspective and examines the role of well-being and flourishing for long-term mental health. In their investigation, several different scales of well-being were used to construct an overall well-being measure. These scales reflected three dimensions of well-being. These were subjective well-being [measured with the Positive and Negative Affect Scale (Watson et al., 1988) and the Satisfaction with Life Scale (Diener et al., 1985)], psychological well-being [measured with the Personal Mastery Scale (Pearlin & Schooler, 1978) and the Connor-Davidson Resilience Scale (CD-RISC) (Connor & Davidson, 2003)], and social well-being [measured with two global scales (distinguishing the positive and negative perspective) of social support from a partner, family and friends (Schuster et al., 1990)].

To construct the overall well-being measure, Burns and colleagues applied Exploratory Factor Analysis (EFA) on the set of above-mentioned well-being measures, using the principal axis method of factor extraction, and derived 1 factor (since only 1 factor reported an eigenvalue >1) that explained 77% of the shared variance. This factor (represented by the factor scores) constituted an overall well-being variable and was subsequently categorized according to the tertiles of the distribution to distinguish (1) languishing respondents (respondents scoring the lowest on the overall well-being variable), (2) respondents with moderate wellbeing, and (3) flourishing respondents (those with highest overall scores in terms of the well-being variable).

There is a lot of value in Burns et al.’s (2022) argumentation that well-being is a multidimensional concept and its complexity warrants measurement with a number of diverse indicators. This conviction is also substantiated by the existence of plethora of multidimensional measures of well-being as well as numerous research on distinguishing subscales to measure various dimensions of the concept (a few examples are Marsh et al., 2020; Ryff & Keyes, 1995; Su et al., 2014; Weziak-Bialowolska et al., 2021). Yet, as authors duly noted, there has been little consistency in how flourishing is defined across existing frameworks. One approach is to define flourishing as the state in which one scores the highest on a well-being scale [the approach adopted by Keyes (2002), Keyes (2005), Keyes and Simoes (2012) and also by Burns et al. (2022), among others], another is to define flourishing as a distinct concept and construct an alternative instrument designed to measure it. In this vein, VanderWeele (2017) defines flourishing as a state of complete well-being which captures these life areas that are universally perceived by people as essential to their well-being. These life areas comprise being healthy, happy, having meaningful and purposeful life, being a good person of excellent moral character and having fulfilling relationships. Alternatively, Diener et al. (2010) focus on psychosocial flourishing that is reflected in one’s self-perceived success in important life areas such as

* Sustainability and Health Initiative (SHINE), Department of Environmental Health, Harvard T. H. Chan School of Public Health, 655 Huntington Ave, G28, Boston, MA, USA.

E-mail address: dowezia@hsph.harvard.edu.

https://doi.org/10.1016/j.ssmmh.2021.100060
Received 19 December 2021; Accepted 30 December 2021
Available online 3 January 2022
2666-5603/ © 2022 The Author. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
relationships, self-esteem, purpose, and optimism. A specifically designed measures of flourishing are proposed in both cases, these are Flourishing Index and Flourishing Scale, respectively.

I share the conviction of Burns et al. (2022) that the use of a range of indicators to capture the diversity of dimensions that underpin well-being constitutes an important prerequisite for its valid measurement [a very informative discussion on comprehensive measurement of well-being with longer and shorter instruments has been recently presented by VanderWeele et al. (2020) with the dissenting voice of Ryff et al. (2020) and the response by VanderWeele et al. (2021)]. The reports of Burns et al. (2022) offered a valuable perspective on the inconsistent terminology, numerous frameworks, operationalizations, and settings in which measurement of flourishing occurs as well as on flourishing instruments proposed. These are certainly important aspects of the complexity inherent to the topic of human flourishing. Despite the fact that the work of Burns and colleagues provides an important contribution in the field of promoting human flourishing and examining its role for mental health, I do not share their view on the flourishing operationalization that has been presented in their study. In my opinion, it remains at odds with the current recommendations of the psychometric literature. I would like to offer two perspectives on this: (1) the way the overall well-being variable was constructed, and (2) the way the flourishing concept was operationalized.

Regarding the former, to ensure valid and reliable measurement, an instrument should be carefully defined and constructed. The procedure comprises both theoretical considerations and psychometric analysis. While there is no doubt that it is important to use a range of indicators to capture the diversity of dimensions that underlie well-being, as Burns et al. (2022) did, and while I deeply appreciate their effort to use secondary data and existing instruments, I also disagree with the approach in which these indicators and instruments are converted into a single score without a more in-depth examination of the validity and reliability of the composite measure (i.e., the overall well-being, in this case). Although Burns and colleagues provided some statistical evidence on one-dimensionality of the overall well-being measure (i.e., they identified only one eigenvalue above 1; 77% of explained variance in the exploratory factor analysis supported by the evidence on its configural, metric, and most likely scalar measurement invariance in terms of age), the consensus in the literature is to conduct a more thorough analysis of the structure and dimensions of any proposed instrument.

One of the suggested procedures to ensure well-defined factor structure is to apply the construct validity approach in which the construct and the instrument are empirically examined using either item-response theory (IRT) or factor analysis (exploratory and confirmatory) with an evaluation based on the goodness-of-fit indices, like Comparative Fit Index (CFI), Tucker-Lewis Index (TLI) and Root Mean Square Error of Approximation (RMSEA). Yet, even such analysis is not considered sufficient. Further analyses are needed to ensure other types of validity. First, they can include evaluation of correlations between the instrument of interest (and also its dimensions) and the socio-economic characteristics of respondents, which allows to position the instrument in a broader conceptual space, ensuring a theoretically and logically coherent pattern of associations between the instrument and external variables. Second, stability of the factor structure, ensured by testing and re-testing the model based on responses from multiple time points and establishing measurement invariance with respect to time, allows to provide valid claims that the functioning of instrument does not change over time. Third, the fact that is invariant across basic demographic characteristics should be also desired as it allows to claim that different sub-populations have similar understanding of the concept and that the instrument can be used to measure well-being overall and across specific groups. Fourth, convergent and discriminant validity should be examined in relation to well-established measures reflecting similar and dissimilar concepts. While I recognize that the third option might have been beyond the scope of the study by Burns and colleagues (their focus was on the life course, which constitutes an entirely understandable reason for their interest in age cohorts and justifies examining the measurement invariance across the age cohorts only), the remaining ones, if addressed, could have provided better foundations for the presented results.

Moreover, validity of a new instrument is not the only concern – its reliability needs to be assessed as well. In this respect it has been agreed that demonstration of satisfactory reliability of an instrument using a reliability measure (the most popular is the Cronbach’s alpha but other measures have been also proposed especially in light of the criticism of this standard approach [for details see Bentler, 2009; Sijtsma, 2009]) and test–retest reliability over time is expected.

The issue worth addressing in this commentary is also the use of parcel-indicators in factor analysis and structural equation modeling. While item-indicators are singular survey items, item parcels are the measures obtained in the process of aggregation (usually sum or mean) of responses to several indicators designed to measure the same construct. There is a discussion whether parcel-indicators can be used in applied science, but overwhelming argument is that more reliable results are obtained when the item-indicators are used (Little et al., 2002; Marsh et al., 2013; Wen et al., 2013). Parcel-indicators, however, can be used but only under very specific conditions. It has been theoretically argued and empirically shown that they can be used only if there is good a priori information supporting the postulated factor structure at the item level. This means that each item-indicator should load on one factor only (i.e., there are no cross-loadings), measurement errors should be uncorrelated and no secondary factors should exist. This further implies that in order to apply parcel-indicators, one should first run item-level factor analysis to examine whether a traditional independent clusters confirmatory factor analysis model (ICM-CDA) fits the data at the item level. Evaluating the correlation matrix (for item-indicators) and checking item level reliability are also important first steps in parcel formation process and should not be omitted even if the goal of a study is only to examine the relations among constructs. These steps are even more important when the aim of the study is scale development (Marsh et al., 2013; Wen et al., 2013). Although Burns et al. (2022) are silent whether they used parcel-indicators or item-indicators in their operationalization of the overall well-being, their report on only one eigenvalue above 1 may be indicative that parcels were used.

I entirely share Marsh et al.’s (2020) view that a combination of different indicators that are aggregated by using sum or mean or factor scores, if ill-defined and not substantiated by a well-established factor structure, impedes our understanding of what precisely is being measured. Therefore, it is of utmost importance to present satisfactory theoretical and statistical evidence supporting psychometric properties of any newly proposed instrument. Despite being cognizant that the scope of analyses necessary to ensure acceptable psychometric properties is extensive, I am missing presentation of all these details in the study of Burns and colleagues (2022). Even though these results would probably take substantial space, these details are always expected to be presented when designing and ‘promoting’ a new instrument.

Whether (i) a newly designed instrument with entirely new items, (ii) an instrument with both a subset of items from already established measure(s) and secondary data (i.e., not from a specifically designed study) such as the instrument by Burns and colleagues, or (iii) an instrument using a subset of items from already established measure(s) and primary data [for example, a new combination of items already used in prior studies was used in the Well-Being Profile (WB-Pro) by Marsh et al. (2020) or in the Flourishing Index by VanderWeele and colleagues (VanderWeele et al., 2013; Westak-Bialowolska et al., 2021)] is proposed, it is usually expected that first, a validation study is published. Only after a successful publication of the validation study, it is expected that studies in which the instrument is applied can be submitted for publication. This two-step process takes time and requires substantial efforts. I wonder whether a more practical approach could be applied. That is all necessary analyses needed to validate a measure could be presented with due diligence as a supplementary material for a primary publication. In such a case, scholars willing to present their research, in
which a new measure is applied but this measure itself does not constitute a focal point of the study, would not have to wait for a validation study to be published first, and could proceed with a submission of their primary research. This strategy could have been perhaps applied in the case of Burns et al. (2022). An open question remains, however, whether journal editors are ready to affirm and reviewers are prepared to positively perceive such a practice.

The operationalization of flourishing calls for a comment, too. Burns et al. (2022) decided to define flourishing respondents as those who scored above the second tertile threshold for the overall well-being variable. Although a similar strategy to define flourishing was adopted in prior studies (see, for example, Keyes (2005) and Keyes and Simoes (2012) who, using an analogy to a mental disorder that can be perceived as a deviation from the typical mental functioning, defined flourishing as an above-average functioning and languishing as a below-average functioning), I believe that to advance studies on flourishing and in particular to make flourishing a concept that may inform public health policy and clinical practice, a sample-independent cut-off points defining flourishers and languishers are needed. Merely scoring the highest (or in the top tertile) does not guarantee flourishing as this is a purely relative identification. In a situation of a strongly positively skewed distribution of an overall well-being variable, there may be substantial group of subjects, whose score – despite being located in the top tertile – would be far from the maximum score, which theoretically should be aimed for. Although the task to establish more objective thresholds is certainly not a trivial task, I think that this is what proponents of human flourishing should aim for. Without such objective thresholds, we cannot expect a clinician to use in her practice an instrument designed to diagnose flourishers/languishers. Therefore, I fully support Burns et al.’s (2022) call for continuing research on the prospective role of well-being for future mental health outcomes. I would only add to this that the focus on flourishing is also of crucial importance.

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References


D. Węziak-Bialowolska

SSM - Mental Health 2 (2022) 100060