ORIGINAL ARTICLE

Comparison of wellbeing structures based on survey responses and social media language: A network analysis

Selim Sametoğlu[1,2] | **Dirk H. M. Pelt**[1,2] |
Johannes C. Eichstaedt[3,4] | **Lyle H. Ungar**[5,6] | **Meike Bartels**[1,2]

[1]Department of Biological Psychology, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

[2]Amsterdam Public Health Research Institute, Amsterdam University Medical Centers, Amsterdam, The Netherlands

[3]Department of Psychology, Stanford University, Stanford, California, USA

[4]Institute for Human-Centered AI, Stanford University, Stanford, California, USA

[5]Department of Computer and Information Science, University of Pennsylvania, Philadelphia, Pennsylvania, USA

[6]Positive Psychology Center, University of Pennsylvania, Philadelphia, Pennsylvania, USA

Correspondence
Selim Sametoğlu, Department of Biological Psychology, Faculty of Behavioural and Movement Sciences, Vrije Universiteit Amsterdam, Van der Boechorststraat 7-9, 1081 BT Amsterdam, The Netherlands.
Email: s.sametoglu@vu.nl

Abstract

Wellbeing is predominantly measured through surveys but is increasingly measured by analysing individuals′ language on social media platforms using social media text mining (SMTM). To investigate whether the structure of wellbeing is similar across both data collection methods, we compared networks derived from survey items and social media language features collected from the same participants. The dataset was split into an independent exploration ($n = 1169$) and a final subset ($n = 1000$). After estimating exploration networks, redundant survey items and language topics were eliminated. Final networks were then estimated using exploratory graph analysis (EGA). The networks of survey items and those from language topics were similar, both consisting of five wellbeing dimensions. The dimensions in the survey- and SMTM-based assessment of wellbeing showed convergent structures congruent with theories of wellbeing. Specific dimensions found in each network reflected the unique aspects of each type of data (survey and social media language). Networks derived from both language features and survey items show similar structures. Survey and

SMTM methods may provide complementary methods to understand differences in human wellbeing.

KEYWORDS

language, natural language processing, psychological networks, social media, wellbeing, well-being

INTRODUCTION

The interest in the concept of wellbeing is increasing, given its relation to various positive outcomes. Higher wellbeing is associated with better finances and social relations, more altruistic behaviours, higher school achievement and better workplace functioning (Chapman & Guven, 2016; James et al., 2019; Kim et al., 2019; Maccagnan et al., 2019; Okabe-Miyamoto & Lyubomirsky, in press; Oswald et al., 2015; Steptoe, 2019; Walsh et al., 2018). Increases in wellbeing is found to be prospectively associated with lower healthcare costs and sickness benefits in randomised nationally representative samples (Santini, Becher, et al., 2021; Santini, Nielsen, et al., 2021) and other studies. If supported by governmental policies, positive wellbeing may boost the socio-economical development of nations, because evidence suggests higher wellbeing levels are associated with lower healthcare costs (Santini et al., 2021; Sears et al., 2013; Shi et al., 2013), as well as job turnover and productivity loss (Sears et al., 2013; Shi et al., 2013).

There is a wide range of conceptualisations and measures for wellbeing. Overall, most wellbeing measures can be categorised under subjective (or *hedonic*) and psychological (or *eudaimonic*) wellbeing (Deci & Ryan, 2008; Ryff, 1989). Subjective wellbeing (SWB), reflecting a hedonic conceptualisation of wellbeing, consists of cognitive and affective evaluation of one's life. The cognitive component of the SWB is captured by life satisfaction, whereas the affective component is measured by (the presence of) positive affect and (the absence of) negative affect (Diener et al., 1985). Psychological wellbeing (PWB; Ryff, 1989), based on an eudaimonic conceptualisation of wellbeing, is defined as positive functioning in life, consisting of positive relations, autonomy, environmental mastery, personal growth, purpose in life and self-acceptance. Despite their unique features, most wellbeing measures at least moderately correlate with each other, suggesting an underlying common, broad wellbeing factor (Bartels & Boomsma, 2009; Baselmans & Bartels, 2018; Disabato et al., 2016; Longo et al., 2016).

The majority of studies investigating wellbeing are based on self-report questionnaires (Dolan et al., 2011; also see p. 154 in Proctor & Tweed, 2016). Although shown to be reliable, use of questionnaires is time consuming and may be prone to biases such as social desirability (Edwards, 1957), recollection bias (see, for instance, Shiffman et al., 1997) or wording effects (Schuman & Presser, 1996). In addition, questionnaire items capture wellbeing over extended time periods by asking individuals to reflect on their past or general life at present (e.g. wellbeing over the last 6 months or 'how happy are you in general?'), providing a static measure of wellbeing. A potential alternative for the assessment of wellbeing is through the analysis of individuals' language expressed on social media over an extended time period.

In 2020, more than 3.6 billion people worldwide used social media platforms (Tankovska, 2020). Social media platforms provide a medium for people to share their thoughts, emotions and behaviours in real time, which can be accessed unobtrusively. As a data-driven method that summarises users' language use, Latent Dirichlet Allocation

(LDA; Blei et al., 2003) allows for identifying naturally occurring word clusters in individuals' language on social media based on a data-driven approach. These word clusters, called LDA topics, include words that have related meanings. For instance, a language topic may include words such as 'tomorrow, excited, nervous', whereas another topic includes words like 'friends, family, thankful'. For each individual, relative frequencies of these topics can be calculated, resulting in topic usage scores. A higher topic usage score indicates that the given person has a higher probability of using the words included in a particular topic compared with individuals with a lower score. Topic usage scores can be thought of as analogous to survey scores and have been utilised previously for the automatic assessment of a wide range of traits based on the language expressed on social media such as personality (Park et al., 2015), depression and mental illness (Eichstaedt et al., 2018; Guntuku et al., 2017). In LDA, different but related words are compiled into topics, whereas in the survey measures, different items are used to assess an overarching construct (e.g. satisfaction with life items aggregated into a satisfaction with life sum score) (for a review, see Eichstaedt et al., 2021).

LDA is one of the various methods available to automatically analyse the language of individuals on social media to assess traits, which collectively can be referred to as social media text mining (SMTM) methods. SMTM methods can provide an unobtrusive and real-time measurement of wellbeing based on the language coming from the social media accounts of individuals. It is known that survey- and SMTM-based wellbeing correlate with each other (meta-analytic estimates of .54, 95% CI [.37, .67] for location-level studies, and .33, 95% CI [.25, .40]) for individual-level assessments (Sametoglu et al., 2022). Such level of convergence obtained between the survey- and SMTM-based wellbeing scores are considered as an indicator of (convergent) validity. However, these correlations are based on aggregate measures of wellbeing (sum of survey items and the predicted wellbeing levels based on all topics features). Thus, at a more fine-grained level, it is still unclear to what extent the wellbeing survey items and social media language features associated with wellbeing are structurally similar. Alternative methods are needed to address these research questions.

To compare wellbeing assessments based on surveys and SMTM in a different way, and to further evaluate the validity of the SMTM method, a network analysis can be employed to examine both the structural and content-based similarities between the two methods. More specifically, by using partial correlation networks (Borsboom & Cramer, 2013) with exploratory graph analysis (EGA; Golino & Epskamp, 2017), the complex relations and data-driven clusters occurring among the wellbeing survey items and SMTM-based wellbeing topic use scores can be mapped out. In addition, both survey- and language-based partial correlation networks can be compared in terms of their similarities with regard to their dimensions and whether their dimensions both align with the existing theoretical conceptualisations of wellbeing. For a more thorough examination, it is also possible to evaluate the most central items in the network. This can be achieved by estimating different network centrality measures (e.g. *closeness*, *betweenness*, *strength*), which can be also compared across both networks, but also with the existing wellbeing literature.

Relatedly, a number of studies has previously applied the network approach to understand the structure of survey-based wellbeing, whereas similar studies are lacking for social media language-based wellbeing measures. The survey-based network studies have provided information on which wellbeing items appear to be most important in terms on the strength centrality metric (average strength of the edges that a node had with the other nodes), in which their results were largely consistent among each other. Stochl et al. (2019) used the 14-item Warwick-Edinburgh Mental Well-being Scale (WEMWBS; Tennant et al., 2007), aimed to

capture both hedonic and eudaimonic aspects of wellbeing, applied on a sample consisting of four UK cohorts. The most central three items were about self-acceptance, self-confidence and cheerfulness. Similarly, a study by Zeng et al. (2019) used network analysis to study wellbeing items related to engagement, perseverance, optimism, connectedness and happiness (see the EPOCH wellbeing survey; Kern, Benson, et al., 2016; Kern et al., 2019) on a Chinese adolescent sample. They found the items with the highest centrality were about being absorbed in the current activity, optimism and cheerfulness. In another study, Van de Weijer et al. (2021) estimated a network on a wide variety of measures including satisfaction with life, subjective happiness, quality of life, flourishing, self-rated health, depressive symptoms, neuroticism and loneliness. The most important three items (based on the strength centrality) were about overall satisfaction with life, feeling unworthy or inferior. Lastly, a study by Giuntoli and Vidotto (2021) applied network analysis on SWB (Diener et al., 1985), flourishing (Diener et al., 2010) and affect (via the Scale of Positive and Negative Experience [SPANE]; Diener et al., 2010) in an Italian adult sample. Their results revealed that the life satisfaction item 'I am satisfied with life' was the most central item in their network. Overall, the mentioned studies have used partial correlation networks to map out the complex structure of wellbeing based on the survey items. The results obtained from these studies primarily highlighted self-acceptance/self-worth as a recurring theme among their most central and influential items (as long as an item related to such a theme was in the study). These previous studies, however, have failed to include more detailed measures of eudaimonic wellbeing such as the Ryff Scales of Psychological Wellbeing (PWBS) (Ryff, 1989). Further, the previous studies have not used network indices on the global state of a network, but rather focused on node-level characteristics. These node-level characteristics such as strength, closeness or betweenness only provide information on the importance of an individual node relative to the total network structures, with the aim for finding the most relevant items for improving wellbeing of individuals. However, when one is interested in the structure of wellbeing as a construct, the network should also be viewed in its entirety, that is, as a system (Borsboom et al., 2021). Inclusion of global network indices can provide information with regard to the extent to which wellbeing networks (as summarised through surveys or other measures) are densely connected, clustered or homogenous, which cannot be otherwise provided through by only looking at the node-level characteristics.

Contrary to the more frequent application of network methodology on the survey-based wellbeing measures, the number of studies applying this method on language is limited. As one of these few examples available, Kjellstrom and Golino (2019) have applied data-driven network analysis to identify the most important health responsibility themes emerging in interview transcripts from different age groups (Kjellstrom & Golino, 2019). Further, to our knowledge, no study has leveraged network methods to assess the structure of wellbeing as reflected by SMTM-based measures. Estimating networks on social media language topics combined with EGA can provide a valuable opportunity to lay out the structure of SMTM-based wellbeing and further allow to find which data-driven wellbeing dimensions exist in such language-based social media data. Importantly, it allows for an investigation of the validity of SMTM-based wellbeing by comparing its network structure with the network of survey-based wellbeing.

THE PRESENT STUDY

The networks estimated using survey responses and social media language can provide insight into the structure of wellbeing as assessed by each method. Additionally, they can

reveal the degree and manner in which SMTM-based wellbeing assessments are similar to existing survey measures. Such a fine-grained comparison (through the networks) between the survey and SMTM-based wellbeing is important as the correlational convergence shown by the existing literature (Sametoglu et al., 2022) is only based on the aggregate measures of wellbeing (sum of survey items and the predicted wellbeing levels based on all topics features) and does not allow for a comparison between the individual elements (items and language topics) that adds up to overall survey- and SMTM-based wellbeing scores. Therefore, the present study leverages network analysis to provide a detailed comparison between the two methods.

Importantly, in the present study, we use both hedonic and eudaimonic measures of wellbeing. This was because the conceptual breadth of the wellbeing survey measures previously used to correlate with SMTM-based wellbeing scores were mostly limited to hedonic measures of wellbeing (i.e. positive affect or life satisfaction) (Sametoglu et al., 2022). However, contemporary theories of wellbeing (Diener et al., 1985; Keyes et al., 2010; Ryff, 1989) suggest that including both hedonic and eudaimonic wellbeing measures ensure a more comprehensive and detailed picture of an individual's wellbeing levels. Therefore, the comparisons between survey and SMTM-based measures for wellbeing should involve both hedonic and eudaimonic aspects of wellbeing. By doing so, we aim to provide a more exhaustive and complete view on the alignment of both survey and SMTM wellbeing assessment methods. We also include global measures in addition to node-level network measures to expand our view on how survey and language networks align with each other.

METHOD

Participants and procedure

The present study was approved by the Institutional Review Board at the University of Pennsylvania (IRB No. 813866). Adult participants in the United States were recruited through Qualtrics. The analyses were based on the 2169 individuals (59% male) who filled out the questionnaire, consented to share their Facebook data, posted at least 500 words on their Facebook statuses (Kern, Park, et al., 2016), passed the attention check items of the survey and had no missing data for the variables we used. In the sample, the mean age was 41.38 ($SD = 15.6$, range $= 18$–84), and the average yearly household was US$51,771.97 ($SD = 85,140.32$). Twenty per cent of the participants had a high school degree, 28% had an uncompleted college education, 11% had a two-year degree from a college, and 19% had a 4-year bachelor's degree. The remaining 2% had reported having either less than a high school diploma (or no schooling), a non-college education after high school, an uncompleted postgraduate education or a postgraduate degree. The sample consisted of 80% Caucasian, 5% Hispanic/Latino, 9% Black/African American, 1% Native American/American Indian, 3% Asian/Pacific Islander and 2% 'other ethnicities'. The total number of statuses was 1,650,709 ranging from 8 to 5086 per individual ($M = 761.05$; $SD = 793.94$). To ensure higher robustness of our results, we randomly divided our data into an *exploration* ($n = 1169$) and *final* subset ($n = 1000$). The exploration subset was used for deciding on redundant survey items and language topics to be excluded from our final networks and for detecting any potential problems such as wording effects in our survey measures. On the contrary, the final subset was only used to estimate final survey- and language-based wellbeing networks. The decision for the number of individuals to be included in each

subset was based on power analyses that were performed prior to any of our analyses (see Supporting Information S1).

Survey measures

The items of the following surveys were included in the exploration network.

Quality of life

To measure global life satisfaction on a scale of 0–10, a single-item measure, the Cantril Ladder was used (Cantril, 1965). The item is: 'Please imagine a ladder with steps numbered from 0 at the bottom to 10 at the top. The top of the ladder represents the best possible life for you and the bottom of the ladder represents the worst possible life for you. On which step of the ladder would you say you personally feel you stand at this time?' Single-item measures for wellbeing (such as Cantril Ladder in this case) are usually found reliable (Lucas & Brent Donnellan, 2012).

Satisfaction with life

The Satisfaction with Life Scale (SWLS; Diener et al., 1985) is a 5-item measure of global life satisfaction. Participants responded to each item on a 7-point Likert scale with 1 meaning *strongly disagree* and 7 indicating *strongly agree*. Two example items are: 'My life is going more or less as I wished' and 'I am satisfied with life'. The internal consistency of SWLS in the present study (based on the full sample) was high (coefficient alpha = .91).

Flourishing

The Flourishing Scale (Diener et al., 2010) is a measure of eudaimonic wellbeing consisting of eight items. Participants responded to each item on a 7-point Likert scale (1 *strongly disagree* and 7 *strongly agree*). Example items are: 'I lead a purposeful and meaningful life' and 'I am engaged and interested in my daily activities'. The internal consistency of this scale (based on the full sample) was .92.

PWB

The Ryff PWBS (Ryff, 1989) was used to measure eudaimonic wellbeing and its six theoretical subcomponents: autonomy, environmental mastery, personal growth, positive relations, purpose in life and self-acceptance. The scale consists of 42 items in total and seven items per dimension. Example items for each of the eudaimonic wellbeing dimensions are as follows: 'I have confidence in my opinions, even if they are contrary to the general consensus (autonomy)', 'In general, I feel I am in charge of the situation in which I live (environmental mastery)', 'I think it is important to have new experiences that challenge how you think about

yourself and the world (personal growth)', 'People would describe me as a giving person, willing to share my time with others (positive relations)', 'Some people wander aimlessly through life, but I am not one of them (purpose in life)', 'When I look at the story of my life, I am pleased with how things have turned out (self-acceptance)'. The internal consistency coefficient (based on the full sample) for the total scale was .94, whereas for the subscales the coefficients ranged between .75 (personal growth) and .88 (self-acceptance).

Social media language measures

Social media wellbeing topic usage scores were based on Facebook data. In the present study, each survey participant provided access to their Facebook status updates. After acquiring their raw Facebook data, all non-English data were filtered and deleted, URLs and mentions (indicated by '@') were replaced with <URL> and <USER> tags, and duplicate Facebook posts were removed. Subsequently, 1-g language features (i.e. single words) were extracted as the unit of analysis for the language features. We deleted words that were used by less than 5% of the participants. In line with Kern, Park, et al. (2016), participants with less than 500 words were excluded from the analyses. Next, topic usage scores were calculated for each individual, based on an existing weighted LDA topic lexicon consisting of 500 language topics and their associated weights for reflecting the wellbeing of individuals (Eichstaedt et al., 2021). LDA was performed via the Mallet package (McCallum, 2002) on a random subset of 5 million Facebook statuses provided in the myPersonality dataset (Stillwell & Kosinski, 2004). This allowed us to calculate 500 different topic scores for each individual in our sample.

Statistical analyses

The preregistered analyses, and the deviations made from this initial plan, can be accessed from the following link: 10.17605/OSF.IO/XPDWR

Topic selection and pruning

LDA topics are often near-duplicates. In network models, having duplicate topics is not desirable. Therefore, we reduced the number of topics to be included in our language-based network in the exploration dataset. To de-duplicate the set of associated topics, the 500 topics were sorted in order of their correlation with each of the nine survey-based wellbeing sum scores (Cantril Ladder, SWL, Flourishing and the six subscales Ryff PWBS). The Benjamini–Hochberg procedure (BH; Benjamini & Hochberg, 1995) was applied on the p-values to control the false discovery rate. For each sorted column of topics (nine in total), a lower-ranking topic was not included if more than 25% of its top 15 words were also contained in the top 15 words of a higher-ranked topic. In this process, we limited the final number of topics in our list to be between 50 and 60 to have a similar number of nodes as in our survey-based wellbeing exploration network. To align the survey and network wellbeing scores, we reverse coded the topic usage probabilities of topics based on their correlations with the wellbeing scales. If a topic's highest correlation with the wellbeing scales was negative, the usage probability of this topic

was subtracted from 1 (i.e. reverse coded) so that a higher topic score reflected a higher level of wellbeing. The correlations between the selected topics and the wellbeing sum scores are provided in Supporting Information S2.

Network estimation

EGA (R package 'EGAnet'; Golino & Epskamp, 2017) was used to (1) estimate a partial correlation network with graphical LASSO (glasso; Friedman et al., 2008) as well as Extended Bayesian Information Criterion (EBIC; Chen & Chen, 2008) and (2) identify the node clusters via the walktrap algorithm (Pons & Latapy, 2006). EGA reveals dimensions underlying a network structure; thus, it is used for identifying factors similar to factor analysis methods (Christensen et al., 2019; Golino et al., 2020; Golino & Demetriou, 2017; Golino & Epskamp, 2017). The use of the glasso algorithm requires the gamma tuning parameter to be set manually; it is often set between 0 and 0.5 (Foygel & Drton, 2010). A value of 0 means more edges are estimated, but those edges could include more spurious ones (explorative), whereas a value of 0.5 means fewer edges will be estimated, but some true edges may be missed (cautious). We favoured a cautious approach by using 0.5 for our tuning parameter to avoid the risk of false positive edges in our networks.[1] To facilitate comparisons, nodes in our survey networks were categorised in colours in accordance with the dimensions they theoretically belong to (e.g. all SWLS items were coloured the same), and the dimensions they were placed by the data-driven EGA algorithm (e.g. all items in EGA dimension 1 the same colour). For the language-based network, the nodes were only categorised in colours based on their data-driven dimensions, because no prior classification of these nodes were present.

Attenuating for wording effects

After estimating our pre-registered survey- and language-based exploration networks, clear negative and positive wording effects emerged for the Ryff PWBS items. This resulted in all positively worded items forming one cluster and negatively worded items forming another, regardless of their item content. Because of this, the data-driven dimensions detected through EGA were not in line with the hypothesised dimensions (except the autonomy subscale). Fifty-six items formed five factors in which four of these factors (comprising a total of 47 items) included either only positive or only negative items indicative of different subscale, whereas the fifth factor with the mixed worded items contained all autonomy items from the PWBS together with two items that belonged to another subscale. These wording effects and difficulties in replicating the originally hypothesised factor structure of Ryff's scales have previously been reported (Abbott et al., 2006; Burns & Machin, 2009; Sirigatti et al., 2009; Springer et al., 2006; Springer & Hauser, 2006; Triado et al., 2007). To attenuate for these wording effects in the survey-based exploration and final networks, we used the 'residualEGA' from the EGAnet package (Garcia-Pardina et al., 2022; for the details of the model, see Maydeu-Olivares & Coffman, 2006). Using this method, we fitted a latent wording/method factor to account for wording effects, after which the EGA dimensions were modelled based on the remaining 'residual correlation matrix'. We also applied *MDSnet* from the *networktools* package that makes the distances between the nodes interpretable (Jones et al., 2018), meaning nodes more related to each other are also depicted closer.

Network trimming

After estimating our survey- and language-based exploration networks, we checked possible redundancies in our item and language topic sets. That is, if two separate nodes have similar correlations with other 'third variables' in a network, they most likely capture the same construct, indicating redundancy. We used the *goldbricker* function from the *networktools* R package (Jones, 2022) to check for redundancy and eliminate redundant items/language topics (i.e. '*network trimming*') in our exploration network before estimating our survey- and language-based final networks. This function calculated the proportion of significantly different correlations for each node pair. The items/language topics were discarded if a node pair had less than 50% of their correlations different (thus indicating high levels of similarity). The resulting list of item/language topics was used to estimate the survey- and language-based networks in our final sample.

Evaluation

Network evaluation

To evaluate the characteristics of the networks as a whole, we used the following three measures: the *global clustering coefficient*, *density*, *small-world-ness* and *average predictability*. The global clustering coefficient (i.e. transitivity) reflects how frequent a single node's neighbour nodes are also connected to each other—or how much the network is 'clustered' (Costantini et al., 2019)—and is calculated by dividing the number of close triads (a group consisting of their nodes that are all connected to each other), over the possible triads (a group of three nodes where each node is not necessarily connected via a direct path). The global clustering coefficient can range between 0 and 1, where 0 reflects none of the triads are closed and 1 indicates all triads are closed. A network with a high global clustering coefficient can be interpreted as highly connected and clustered, whereas a low global clustering coefficient is an indication that the network consists of a high number of weak ties. Density is a measure of overall connectedness in network, and it is equal to the number of edges present relative to the number of theoretically possible edges. Density can take values between 0 and 1, the estimates closer to 0 reflecting more sparsity and the estimates closer to 1 indicating a more complex (i.e. 'connected') network structure. Small-world-ness is a situation where global clustering coefficient is high and average path lengths are short (Watts & Strogatz, 1998). Values between 1 and 3 are considered as borderline where values higher than 3 indicate such small-world-ness (Humphries & Gurney, 2008). Average predictability in a network shows, on average, how much the nodes across the network can be predicted by neighbouring nodes (Haslbeck & Fried, 2017). Higher average predictability can be considered as more homogeneity within a network, whereas lower values can be thought as a sign of higher homogeneity (or interchangeability) among the network nodes.

Network node evaluation

To evaluate the importance of each node within our networks, we used the three commonly used centrality measures: *closeness*, *betweenness* and *strength* (Isvoranu et al., 2022). Closeness

shows how strongly a node is indirectly connected to other nodes in a network (calculated by taking the inverse of the sum of distances from one node to all other nodes). Betweenness of a node indicates if particular node has a key role in connecting other nodes and is based on how often one node is in the shortest paths between other nodes. Strength reflects how strongly a node is directly connected to other nodes (calculated by taking the sum of absolute edge weights connected to each node). *Strength* can only be used effectively if the edge weights are positive, as this measure does not differentiate between positive and negative edge weights, which may lead to these values cancelling each other out. Therefore, we also include expected influence, which also assesses strength but considers both positive and negative edges in terms of calculating the strength value.

The network links (or 'edges') in the present study represent statistically significant partial correlations between two nodes while taking into account (i.e. 'controlling for') all other variables in the network. Each network edge has a weight parameter that reflects the strength of such an association between two nodes (see Borsboom et al., 2021 for a more detailed description). All of the global network measures (i.e. global clustering coefficient, density and small-world-ness, as well as the node-specific network measures, i.e. closeness, betweenness, strength, expected influence) are based on the partial correlations between nodes in the network.

Network performance

To assess the accuracy of the edge weights, we used bootstrapping to provide confidence intervals (CIs at alpha = .5) (Bollen & Stine, 1992). For the survey-based final network, we performed parametric bootstrapping by inputting the estimated residualised network structure/weights, whereas for the language-based final networks, we applied nonparametric bootstrapping where portions of the raw data were resampled in each iteration. Nonparametric bootstrapping is commonly used when the underlying distribution of the data is unknown or non-normal, such as in language data ('Zipf's law'; Newman, 2005).

To assess the stability of centrality indices, we used case-dropping bootstrapping method (Epskamp et al., 2018). In each iteration, a bigger percentage of cases/individuals are dropped from the sample, and the centrality measures are re-calculated. The whole process results with a centrality stability coefficient (CS coefficient) for each of the centrality measures (i.e. betweenness, closeness and strength) (Epskamp et al., 2018). The CS coefficient reflects the minimum correlation between the centrality measures based on the full data and the centrality measures based on the subsets of the data. The minimum value for CS coefficient is .25 to signify the stability of the estimates (Epskamp et al., 2018).

Network comparisons

We compared the survey-based and social media language-based final networks on several characteristics: the number of dimensions and edges identified, the value of general network indices (i.e. global clustering coefficient, density and small-world-ness) and the semantic content of each dimension identified in both networks. We also presented and compared the most and the least important nodes in each of the networks based on their centrality scores (betweenness, closeness, strength and expected influence).

RESULTS

Network estimation and trimming

Our survey-based exploration network resulted in 484 positive and nine negative (1.8%) edges (493 in total), whereas our SM language-based exploration network consisted of 299 positive and 90 negative (23.7%) edges (379 in total). The EGA algorithm identified five data-driven clusters in the survey network and nine in the language-based explorative network (see Supporting Information S3–S5 for the related network graphs). The network trimming identified 14 out of 56 items and five out of 52 language topics as redundant (see Supporting Information S6–S9 for the list of items and language topics, respectively).

The survey-based final network using the remaining 42 items included 321 edges [315 positive and six negative (1.9%)]. In Figures 1 and 2, the survey-based final network is presented in two ways: (1) with nodes colour-coded in accordance with their theoretical dimensions and (2) with nodes colour-coded in accordance with their data-driven dimensions. The data-driven approach (based on EGA) suggested that items from nine theoretically defined dimensions can be summarised in five dimensions as depicted (see Table 1 for the list of these items and their related dimensions). The language-based network (as depicted in Figure 3) using the remaining 47 language topics included 325 edges [265 positive and 60 negative (18.5%)]. The data-driven
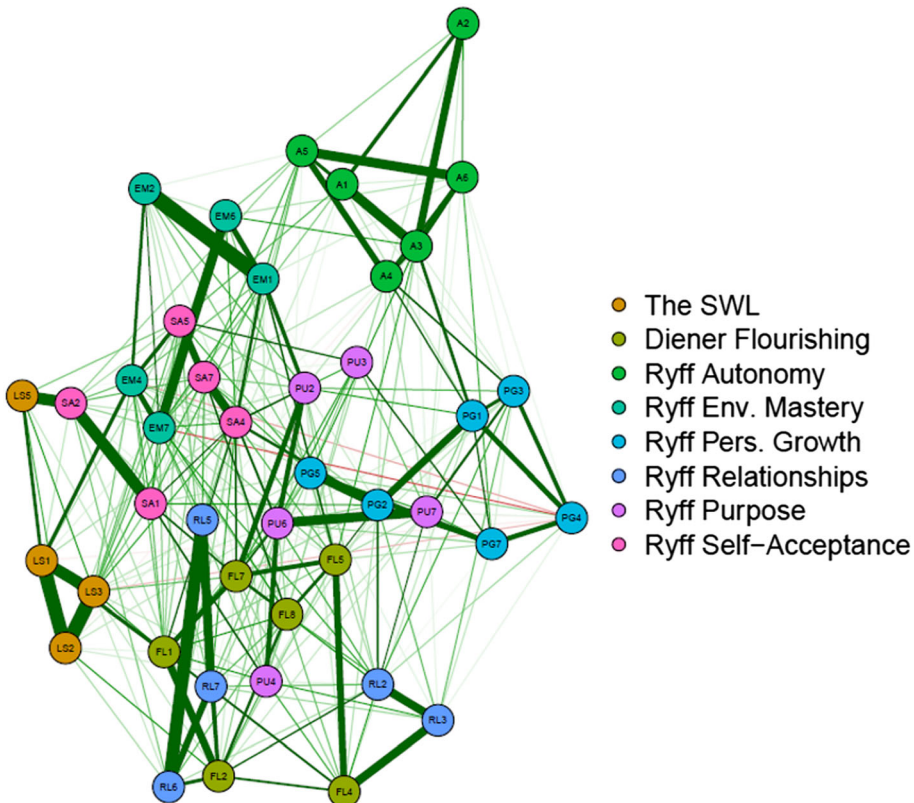

FIGURE 1 Residualised survey-based final network grouped by theory-based dimensions.
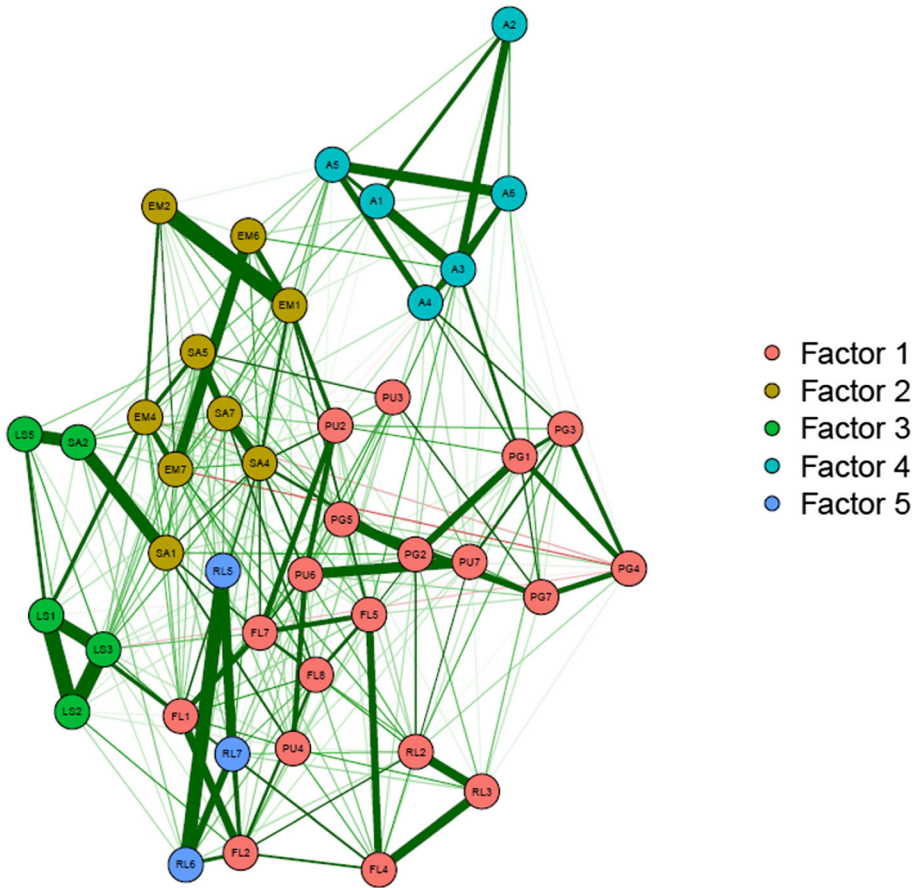
FIGURE 2   Residualised survey-based final network grouped by data-driven dimensions.

approach summarised these topics in five clusters (see Table 2 for list of these topics and dimensions).

Network evaluation

The global network measures for the survey and language-based final networks were as follows: The density values were 0.37 and 0.30, for the survey and language-based final networks, respectively; thus, both networks were sparser than complex (0.50 is the perfect balance), although the language network was sparser. The average predictabilities were 0.51 ($SD = 0.14$) and 0.66 ($SD = 0.15$) for survey and language networks, respectively. This indicated that, on average, an individual node was able to be predicted by other nodes better in the language-based final network than the survey-based final network. The small-world-ness scores were 1.22 and 1.26, indicating borderline values for both networks (Humphries & Gurney, 2008). The global clustering coefficients were 0.50 and 0.37 for the survey- and the language-based final network. These indicated average levels of clustering in the survey and less clustered structure for the language-based final network.

TABLE 1 'Data-driven' (based on the exploratory graph analysis) and theoretical dimensions in survey-based final network.

Item	Dimension	Theoretical dimension	Content
FL1	1	Flourishing	Has a purposeful life
FL2	1	Flourishing	Has good social relationships
FL4	1	Flourishing	Is making others happy
FL5	1	Flourishing	Is capable
FL7	1	Flourishing	Is optimistic
FL8	1	Flourishing	Is respected
PG1	1	Personal growth	Open to new exp.
PG2	1	Personal growth	Self-development
PG3	1	Personal growth	Interested in pers gr.
PG4	1	Personal growth	Interested in growth
PG5	1	Personal growth	Thinks he improved him/herself
PG7	1	Personal growth	Believes one can grow and change
RL2	1	Relationships	Family/friends
RL3	1	Relationships	Thinks he/she is a giving person
PU2	1	Purpose	Is a future planner/worker
PU3	1	Purpose	Is not avoiding the future
PU4	1	Purpose	Thinks there are things to do in the future
PU6	1	Purpose	Knows what to do in the future
PU7	1	Purpose	Still setting goals/motivated
EM1	2	Mastery	Is responsible
EM2	2	Mastery	Is responsible
EM4	2	Mastery	Is responsible
EM6	2	Mastery	Her responsibilities do not overwhelm her
EM7	2	Mastery	Have no difficulty in arranging her life according to his/her standards
SA1	2	Self-acceptance	Accepting the past
SA4	2	Self-acceptance	Self-general confidence
SA5	2	Self-acceptance	Thinks others don't have better life
SA7	2	Self-acceptance	Thinks him/herself is not worse than others in liking him/herself
LS1	3	Subjective wellbeing	Has a life close to his/her ideals
LS2	3	Subjective wellbeing	His/her life conditions are excellent
LS3	3	Subjective wellbeing	Satisfied with life
LS5	3	Subjective wellbeing	Got what he/she wanted from life
SA2	3	Self-acceptance	In general okay with his/her past

(Continues)

TABLE 1 (Continued)

| Item	Dimension	Theoretical dimension	Content
A1	4	Autonomy	Not afraid to tell his/her opinions that are contrary to other people's
A2	4	Autonomy	Not being influenced
A3	4	Autonomy	Has confidence in opinions
A4	4	Autonomy	Thinks his/her own appreciation is more important than the others'
A5	4	Autonomy	Not worrying what others think
A6	4	Autonomy	Not changing his/her mind because of the significant others
RL5	5	Relationships	Not lonely
RL6	5	Relationships	Not lonely
RL7	5	Relationships	Thinks others are not less lonely
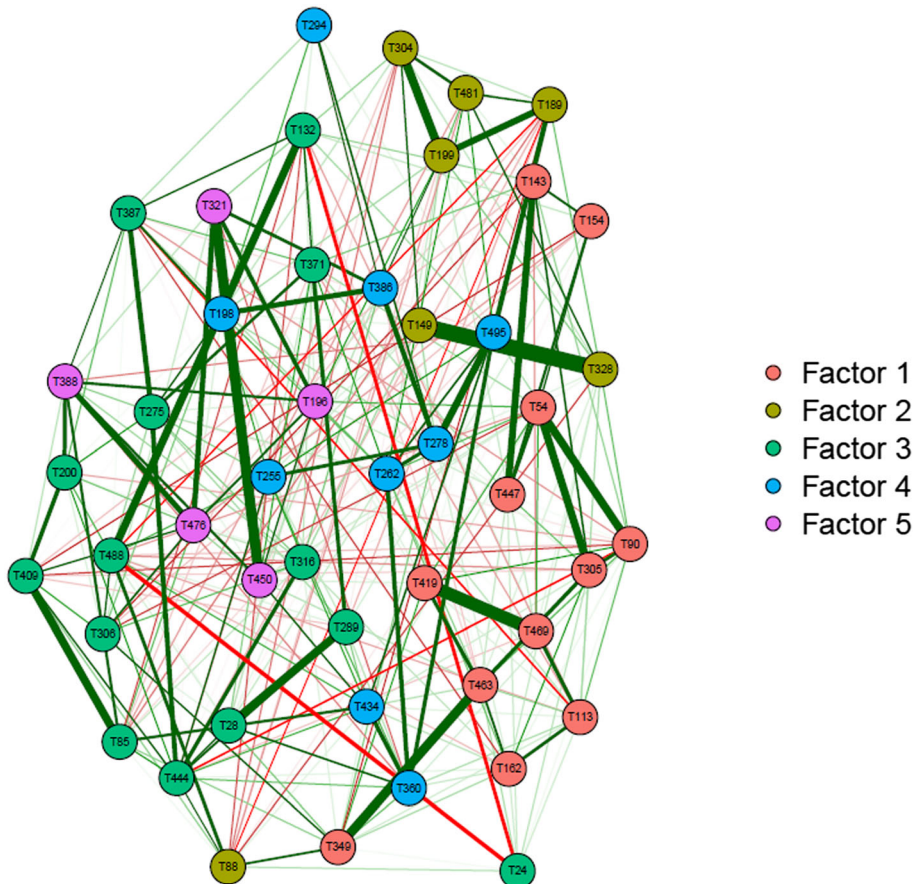
FIGURE 3 Language-based final network grouped by data-driven dimensions.

TABLE 2 Topic clusters identified through EGA on final network.

Topic	Dimension	Words (topic content)
T54	1	human,nature,freedom,beings,power,happiness,fear,humanity,evil, religion,humans,compassion,society,race,lama
T90	1	ignorance,opinion,fact,character,bliss,lack,opinions,personality,respect, judge,ignorant,honest,intelligent,reputation,form
T113	1	word,english,language,speak,french,spanish,words,german,accent,learn, japanese,means,chinese,meaning,british
T143	1	positive,thoughts,mind,negative,thinking,energy,focus,emotions,feelings, attitude,mixed,control,situation,calm,optimistic
T154	1	learn,learned,lesson,teach,lessons,learning,experience,taught,important, learnt,mistakes,boobs,everyday,valuable,teaches
T162	1	read,write,writing,story,wrote,book,reading,letter,poem,written,notes, poetry,note,stories,block
T305	1	reality,human,knowledge,experience,universe,nature,action,sense,create, meaning,simply,exist,existence,logic,imagination
T349	1	college,university,meeting,student,students,accepted,program,library, campus,center,club,conference,letter,event,national
T419	1	government,country,rights,war,america,law,american,state,public, obama,system,states,president,nation,political
T447	1	success,failure,happiness,goal,achieve,goals,reach,key,fail,succeed,ralph, emerson,waldo,successful,greatest
T463	1	business,skills,management,social,career,information,research,company, personal,system,training,major,professional,marketing,medical
T469	1	freedom,american,religion,christian,history,america,country,agree,hitler, religious,christians,flag,russia,jewish,african
T88	2	trip,road,camping,field,planning,adventure,trips,packing,vacation,dum, plan,mission,travel,mini,zoo
T189–f	2	school,tomorrow,starts,monday,summer,classes,forward,starting,tuesday, wednesday,vacation,tommorow,cancelled,orientation,closed
T149–f	2	wait,can't,tomorrow,til,ahhhh,nevermind,tommorow,arsed,tamed, ahhhhhhh,snowboarding,agian,nvm,thrusday,sleeeeeep
T199–f	2	tomorrow,excited,nervous,luck,starting,interview,kinda,scared,anxious, surgery,worried,nerves,excitement,audition,tommorow
T304–f	2	excited,super,sooo,soooo,stoked,pumped,duper,sooooo,uber,psyched, sooooooo,soooooo,prom,junior,mega
T328–f	2	wait,can't,till,weekend,friday,saturday,til,sunday,cnt,2morrow,tomarrow, tommorow,hurry,camping,untill
T481–f	2	woot,yay,whoot,haircut,woohoo,tattoo,braces,wooot,yippie,yippee, payday,babysitting,tomorow,gots,flags
T24–f	3	mom,dad,yrs,mum,years,hero,proud,heaven,angel,blessed,18,16,25,age, annoying
T28–f	3	damn,hell,god,bloody,fucking,dam,dammit,wtf,pissed,damned,fine, damnit,>.<,shut,curse

(Continues)

TABLE 2 (Continued)

Topic	Dimension	Words (topic content)
T85–f	3	don't,care,forget,worry,anymore,understand,talk,bother,judge,expect, honestly,pretend,they'd,they'll,maya
T132	3	friends,family,thankful,wonderful,blessed,grateful,loving,support,lucky, husband,share,helped,caring,supportive,truely
T200–f	3	wrong,what's,there's,whats,point,prove,happening,figure,they're,worse, wondering,difference,happened,supposed,somethings
T275	3	crazy,stupid,insane,normal,smart,dumb,driving,mad,they're,completely, drive,weird,sane,ridiculous,nuts
T289	3	fucking,suck,fuck,dick,balls,ass,bum,sucks,cock,asshole,bitch,fat,fuckin, bastard,douche
T306–f	3	inside,deep,breath,pain,breathe,hide,alive,slowly,holding,dying,skin, control,breathing,scream,breaking
T316	3	stop,can't,thinking,talking,crying,laughing,smiling,loving,calling,quit, stopped,raining,believing,dreaming,worrying
T371	3	guy,girl,he's,dude,random,hes,talking,creepy,chick,gay,named,asks,kid, walks,asian
T387	3	weird,kinda,strange,reason,funny,odd,kind,random,interesting,o.o,hmm, sudden,urge,sounds,finds
T409	3	doesn't,don't,matter,understand,anymore,care,matters,fact,doesnt, realize,exist,trust,important,explain,count
T444–f	3	people,stop,talk,stupid,grow,talking,act,rude,they're,shut,annoying, drama,immature,ignorant,piss
T488	3	family,great,wonderful,dinner,visit,reunion,spent,visiting,cousins, parents,enjoyed,hanging,aunt,evening,spending
T198–f	4	confused,feeling,angry,sad,kinda,mood,upset,depressed,pissed,frustrated, mad,annoyed,lonely,slightly,stressed
T255–f	4	hate,hates,mondays,ugh,suck,passion,hating,absolutely,haters,fact, tuesdays,dislike,allergies,-_-,sundays
T262–f	4	bad,feel,didnt,wow,:/,kinda,wasnt,sucked,felt,totally,forgot,tho,wasn't, worse,=(
T278–f	4	sick,:(,feel,hate,ugh,feeling,sucks,:/,crap,=(,hates,feelin,bein,fml,stomach
T294–f	4	headache,major,worst,horrible,ugh,terrible,attack,boredom,migraine, massive,hangover,panic,sigh,worse,rid
T360–f	4	im,gonna,ill,wont,guess,gunna,ive,didnt,gona,iam,whos,ima,havent, prolly,imma
T386–f	4	tired,i'm,sick,exhausted,mentally,physically,stressed,emotionally, drained,soooo,sore,freaking,completely,extremely,sooooo
T434–f	4	wanna,don't,dont,talk,anymore,cry,scream,billionaire,badly,cover,wana, dun,magazine,fricking,oprah
T495–f	4	school,dont,wanna,:/,ugh,doesnt,sucks,skool,tommorow,tomorow,=/,d:, idk,ughh,bleh

TABLE 2 (Continued)

Topic	Dimension	Words (topic content)
T196–f	5	i'm,i'll,guess,who's,heh,alright,it'll,we'll,eh,suppose,dunno,bet,awhile, eventually,admit
T321–f	5	i'm,afraid,glad,scared,alive,dying,supposed,losing,standing,changing, addicted,fallin,letting,barely,worried
T388–f	5	it's,that's,fine,there's,fault,blame,isn't,line,mine,perfectly,fit,ain't,fair, fucked,shame
T450–f	5	i'm,bitch,mad,addicted,trouble,lucky,drama,ahead,queen,fat,calling,fake, nerd,dressed,freak
T476–f	5	i've,thinking,waste,figure,regret,settle,imma,wasting,else's,trapped,fed, apologize,needing,damned,undone

Note: –f = probability of use for topic was reversed.

Network node-level evaluation

All of the estimated centrality indices are provided in Figures 4 and 5 for the survey- and language-based final networks. The items with highest strength, indicating the highest connectedness to other nodes in the survey network, were SA4 ('In general, I feel confident and positive about myself'), LS3 ('I am satisfied with my life'), A3 ('I have confidence in my opinions even if they are contrary to the general consensus') and EM7 (inverse of 'I have difficulty arranging my life in a way that is satisfying to me') (2.08, 1.53, 1.52, 1.36 SD above the mean, respectively). The topics with highest strength, thus indicating the highest connectedness to other nodes in the social media based final network, were T444 (inverse of 'being angry to or annoyed by other people') and T54 (was largely about 'nature, compassion, and other positive feelings') (2.13, 1.20 SD above the mean, respectively).

In the survey network, the nodes PG2 ('I have the sense that I have developed a lot as a person over time'), A3 ('I have confidence in my opinions even if they are contrary to the general consensus'), SA4 ('In general, I feel confident and positive about myself') and PG1 ('I think it is important to have new experiences that challenge how you think about the world') showed the highest betweenness, meaning that they connected the other nodes to each other the most (2.77, 2.52, 1.82, 1.63 SD above the mean, respectively). In the language network, the topics with highest betweenness were T444 and T54, thus the same as the topic that showed the highest connectedness (3.04, 2.52 SD above the mean, respectively).

The items with highest closeness, meaning the nodes that had the shortest paths to rest of the nodes in the network, were items SA4 ('In general, I feel confident and positive about myself'), FL7 ('I am optimistic about my future') and FL5 ('I am competent and capable in the activities that are important to me') for the survey-based final network (2.31, 1.71, 1.50 SD above the mean, respectively) and topic T444 (i.e. not being angry or annoyed with others) for the social media language-based final network (2.14 SD above the mean, respectively).

The network nodes A2 ('My decisions are not usually influenced by what everyone else is doing') and PU3 ('I tend to focus on the present, because the future nearly always brings me problems') had consistently the lowest centrality scores, hence indicated lower relative importance, in the survey-based final network. For the social media language-based final network, T24 (lack or inverse of commemorating a loved one) and T294 (largely about lack of having headache, migraine and hangover) showed the lowest centrality scores.
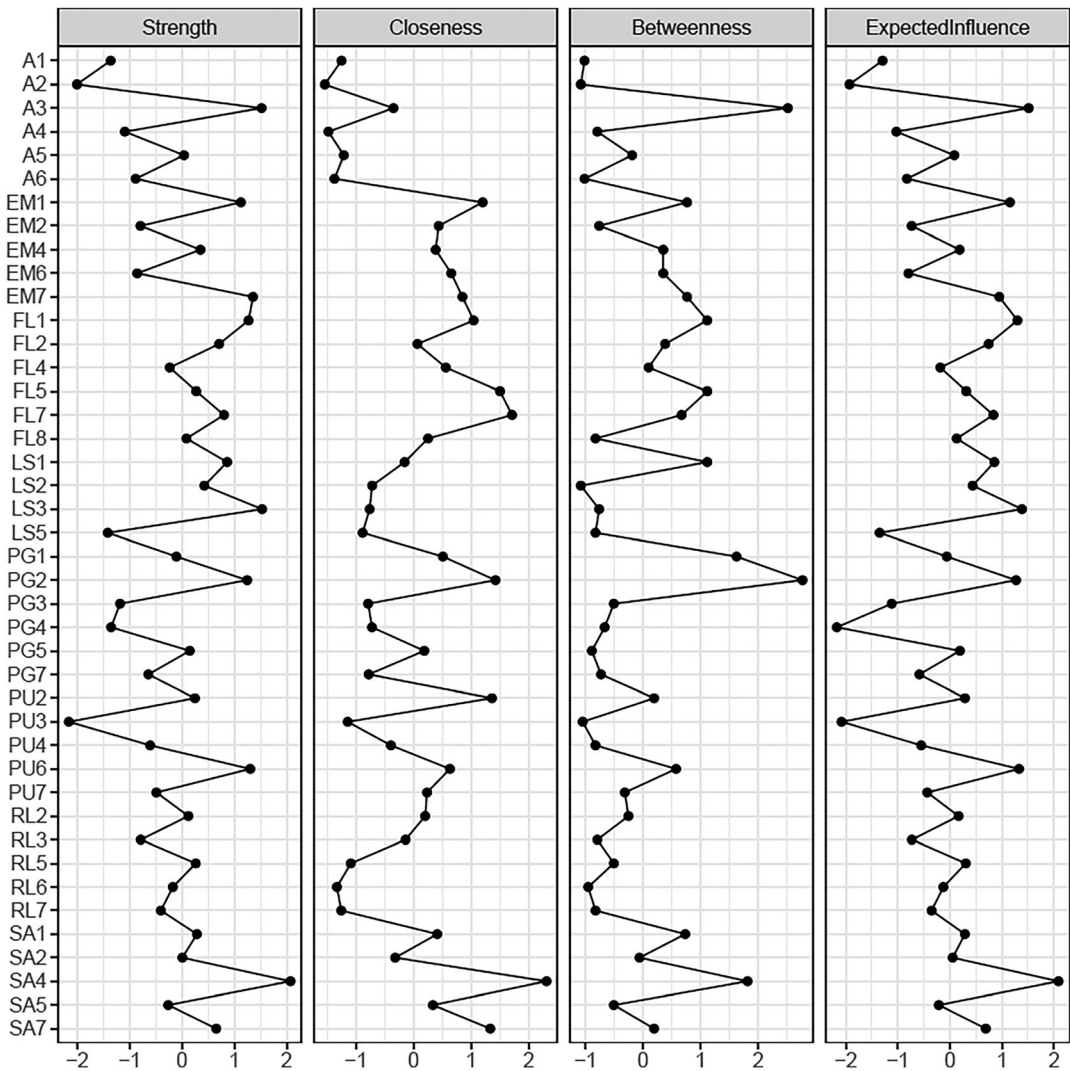
FIGURE 4 Standardised centrality indices for each node in the survey-based network.

Network performance

The edge weight accuracies (i.e. the confidence intervals) for survey-based and language-based final network are provided in Supporting Information S10 and S11, respectively. Visual inspection of these figures showed that edge weights estimated from the sample at hand largely corresponded with the bootstrapped weights and their confidence intervals. As a note, this observation was not true for the smaller edge weights; however, this is an artefact of our regularisation, which sets smaller weights to zero to obtain a better model fit.

We were unable to estimate the centrality stabilities for our residualised survey-based final network (which could have been performed if we did not control for the strong wording effects in our initial network estimation). The centrality stability estimates for the language-based final network, that is, betweenness, closeness, strength and expected influence, were 0.44, 0.36, 0.67
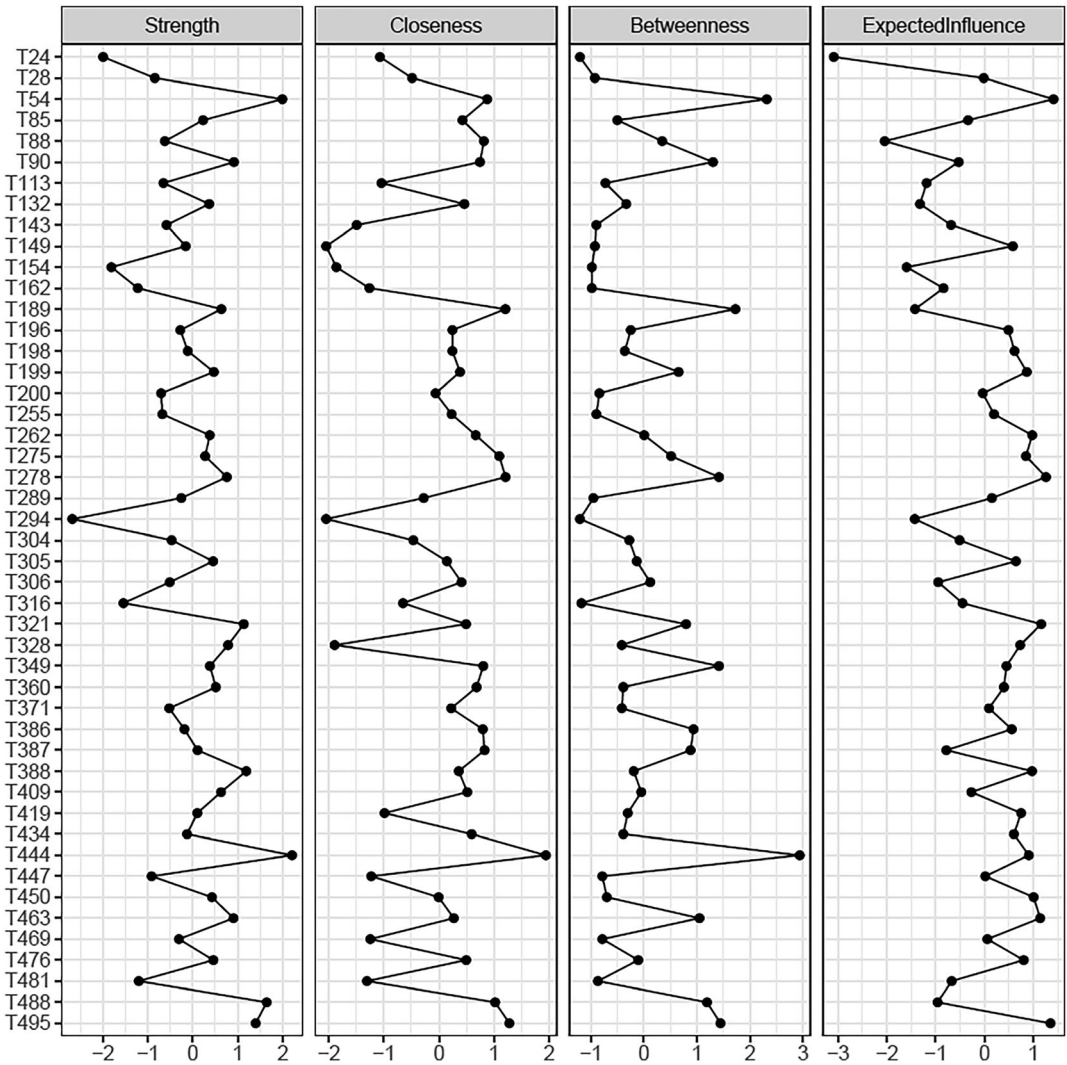
FIGURE 5 Standardised centrality indices for each node in the social media language-based network.

and 0.75, respectively. These values were higher than the minimum value of 0.25, as suggested by previous work (Epskamp et al., 2018), indicating the stability of each of the centrality values we used. Supporting Information S12 and S13 provides a detailed insight into how each of these values had changed at the end of each iteration for survey-based and language-based final networks, respectively.

DISCUSSION

Most studies on wellbeing have been based on survey measures (Dolan et al., 2011; Proctor & Tweed, 2016). Alternatively, individuals' language on social media can be automatically analysed to assess their wellbeing levels as well through SMTM methods. To examine the value

of using SMTM methods to assess wellbeing, we compared the networks based on survey items and language topics (obtained from social media). We first estimated networks to filter out redundant survey items and language topics. After this, final networks based on the remaining items and the language topics were estimated. The final survey-based network showed five wellbeing dimensions, and at the node level, the most important item was about self-acceptance. The final language network was also resulted in five wellbeing related dimensions, and the most important two network nodes were the language topics referring to having less probability for using offensive words towards other people and talking more on the themes related to happiness, compassion and philosophy. Overall, it appeared that both survey and language topic networks conveyed similar information on wellbeing with some differences. Below, we will first compare and interpret the data-driven dimensions found in each network, node-level results and finally the results concerning the edges in both networks. After this, we will provide the present study's limitations, implications and finally the conclusion.

As revealed by our analyses, both survey and language network nodes were summarised in the same number of data-driven clusters. These clusters were similar in terms of their contents and largely mapped onto one another. For instance, in both networks, clear eudaimonic wellbeing, self-acceptance and social support-related dimensions were identified. The dimensions in both networks were theoretically meaningful and reflected constructs found in the wellbeing literature such as eudaimonic wellbeing (Deci & Ryan, 2008; Diener et al., 2010; Ryff, 1989), social support/loneliness (reversed) (Thomas et al., 2017; VanderWeele et al., 2012), self-acceptance/compassion (Deci & Ryan, 2008; Zessin et al., 2015) and satisfaction with life (Diener et al., 1985). Therefore, both networks informed about wellbeing in a similar manner.

Although most of the dimensions in both networks aligned with one another, some of the identified dimensions were unique in each network. First, a life-satisfaction dimension was found in the survey network but not found in the language network. These results may be interpreted as suggesting that humans do not spontaneously engage in evaluating their life and expressing it on social media through language, at least not as much as they do in surveys. Assessing life satisfaction requires the active effort of the individual to make a cognitive judgement of their lives (Diener et al., 1985), and in the survey measures for life satisfaction, the respondents are explicitly asked for making this judgement unlike the SMTM-based assessment of wellbeing. Second, a dimension related to being respectful/kind to others (as indicated by having a lower probability for using negatively valenced words particularly about other individuals) was identified only in the language network, and not in the survey-based network. It appears that social media language partially captures personality dimensions such as agreeableness that is relevant to higher wellbeing (Anglim et al., 2020), but not directly incorporated in wellbeing questionnaires. This is important because survey measures for wellbeing only focus on assessing wellbeing constructs such as happiness or life satisfaction. Third, the presence of a unique dimension in the language-based wellbeing network that reflected not being verbally offensive to others is in line with the meta-analyses indicating a positive association between kindness and higher levels of wellbeing (Hui et al., 2020). The reason for such a dimension to appear only in our language network (but not in our survey network) might lie in the interactive nature of social media platforms that the survey measures do not provide. To be more precise, on social media platforms, individuals may publicly share (or only to a selected group of individuals, i.e. followers or friends) their positive and negative thoughts about a person, group or an entity and also directly engage in conversations with other people. Such expressions may sometimes take forms of antisocial behaviours, which are sometimes known as cyberbullying or

hate speech (ElSherief et al., 2018). Overall, finding unique dimensions shows us that survey and language of social media can inform on different aspects of wellbeing (life satisfaction by surveys, interpersonal communication skills by social media language) and therefore can complement each other to provide a fuller picture of a person's wellbeing.

Concerning our node-level results, the most important nodes in each network were different, but they all aligned with the existing evidence available. In our survey-based final network, 'In general, I feel confident and positive about myself' (SA4) was the most central node (item) as supported by our visual inspection revealing this item had the most frequent links/edges with other items, highlighting its function as a general hub within the overall wellbeing structure. On the other hand, in our language-based network, the most important nodes were about having a lower probability for using aggressive words towards others (based on T444—inversed) and having a higher probability for talking about philosophical matters relating to humans, nature, happiness and compassion (based on T54). The result of the survey network aligned with the most survey-based wellbeing network studies showing self-acceptance (Giuntoli & Vidotto, 2021; Stochl et al., 2019) and self-worth (Van de Weijer et al., 2021) to be among the most central items. In a study by Zeng et al. (2019), no self-esteem or self-worth items were included in their network, thus resulting in items related to activity, optimism and cheerfulness being most important. With regards to our language-based network, we were unable to make any comparison with existing studies applying network analysis to social media language of wellbeing. This was because, to our knowledge, our study is the first one to apply such a method. Nevertheless, the most central topics of our language network were in line with two meta-analyses using survey data indicating both kindness (Hui et al., 2020) and self-compassion (Zessin et al., 2015) to be positively correlated with higher levels of wellbeing.

The differences in most central items between the two networks can be due to the inherent differences between how survey and social media language data are collected and wellbeing scores are created. In surveys, individuals may feel more comfortable to indicate that they feel positive about themselves (as indicated by the most important/central item in the survey network), as the communication is limited between the anonymous participant and the researcher, reassuring individuals to reveal direct information at such a level. On the other hand, on social media, expressing such positive thoughts may be considered as narcissistic as the individuals share their thoughts in the presence of other social media users. Nonetheless, individuals may still express high levels of wellbeing without referring to how precisely they feel positive about themselves, but instead do this through talking more frequently on topics such as happiness, compassion or nature over social media (as indicated by the most central language topics in our study).

With respect to the edges between the nodes in each network, we found that the language network had more homogeneously occurring edges compared with the survey network. Such higher homogeneity might be due to the interchangeable nature of language elements (i.e. words/topics) in daily language use: The same words can be used for expressing different ideas and thoughts. On the contrary, survey items are written to measure a specific construct (for instance, autonomy), thus decreasing the chance for using the same item for measuring another construct (for instance, self-acceptance)—which can explain finding a more clustered/heterogeneous network structure observed in our language networks based on social media language compared with the survey one. This result particularly confirms the more unstructured and versatile nature of social media language data compared with survey measures specifically developed to assess particular wellbeing constructs and dimensions.

Limitations

The present study has several limitations. The sample in the present study was limited to individuals who live in the United States and were paid by the Qualtrics platform to compensate for their participation. Future studies should seek to extend our results in different, more ecologically valid samples. In our study, some of the widely applied centrality measures in psychological networks, which have been recently questioned for their suitability in psychological research, were also used in our study (i.e. betweenness and closeness) (Bringmann et al., 2019). Therefore, the results particularly based on betweenness and closeness should be interpreted with caution and are subject for confirmation by future studies.

The results were based on cross-sectional data and between subject-level analyses. The cross-sectional nature of the data limits any inference to be made regarding the direction of the associations found between the network nodes and the centrality measures (e.g. strength) based on these links. For instance, our finding that self-acceptance is the most central item in our survey-based wellbeing network can be interpreted in multiple ways: It may mean that, compared with other network items, any increase in positive self-view may lead to increases in other wellbeing domains (e.g. personal growth, life satisfaction). It may also mean that any changes in more peripheral nodes (as opposed to being more central) may manifest themselves the easiest on positive self-view levels.

As an important remark about our approach on using these centrality measures, we interpreted results obtained from multiple centrality metrics together to highlight the importance of a single node, instead of interpreting these centrality measures in isolation. This was also because the most central nodes in our networks mostly had the highest centrality scores based on each centrality metric (such as the survey item SA4).

As another potential limitation, we have detected some biases associated with negative wording in our survey responses, that is, items tended to cluster with each other solely based on whether they were phrased negatively or positively. This is highly in line with previous factor analytic studies reporting such difficulties in term of replicating the original six-factor structure of Ryff PWBS (Abbott et al., 2006; Burns & Machin, 2009; Fernandes et al., 2010; Hsu et al., 2017; Sirigatti et al., 2009; Springer & Hauser, 2006; Springer et al., 2006; Triado et al., 2007; for an overview, see Henn et al., 2016). It is also known that survey-based data may suffer from certain biases such as social desirability or recollection bias (Edwards, 1957; Shiffman et al., 1997) or wording effects (Schuman & Presser, 1996) as also found in our study. We have applied specific methods to solve wording-related problems in partial correlation networks (residualEGA; Garcia-Pardina et al., 2022), which appeared to remedy these problems effectively. In addition, although we have not reported similar type of problems in our social media language networks, existing evidence warns about potential biases to be found in social media data as well. For instance, some studies have shown individuals from certain demographics (e.g. sex, educational level, racial background) can be under- or overrepresented in different social media platforms (Hargittai, 2020; Hargittai et al., 2018), which may bias any inferences made at population level based on such data. Kern, Park et al. (2016) mentioned that even though desirability biases are also present in the language of social media, the rank order among individuals remains the same: For instance, even though an introverted individual may try to look a bit more extraverted, on average, they post way less about parties and more on topics like reading compared with the more extraverted individuals. Yet, further investigation is necessary to determine the impact of social desirability on SMTM-based evaluations of wellbeing.

Implications

In general, our findings suggest that both survey-based and social media language-based networks of wellbeing are similar in their ability to provide information about the wellbeing of individuals in line with established theoretical perspectives on wellbeing. At the same time, the differences spotted in each of these networks suggest that the two methods can provide unique information, thus potentially complementing each other. For instance, the topics that happier individuals would like to talk about (e.g. nature, philosophy, education) and how they interact with other individuals (e.g. using fewer offensive language towards others) can be better obtained from social media language data compared with the survey-based measurements of wellbeing. On the other hand, survey-based assessment of wellbeing may better inform about the degree that positive self-view that a person has. Leveraging the information from both approaches used for assessing wellbeing may provide incremental validity over using a single type of a measure.

Our results in terms of finding more homogenous edges in language-based wellbeing networks compared with the survey ones has reflected the fundamental differences in how each data (survey and language data) have been created. The social media language data come from the social media profiles of people, in which the individuals express themselves about topics that are interesting to them and share with others; thus, the data are not limited to a specific theme (i.e. unstructured). On the other hand, survey-based measurement of wellbeing concerns the use of data from individuals in which participants choose their responses from a pre-determined set of answers (e.g. Likert scale format) to also already determined set of questions. Given that social media language data are not limited to a specific topic or construct, it can be leveraged for assessing different constructs (e.g. wellbeing, depression, post-traumatic stress disorder, schizophrenia), whereas surveys allow for assessing a finite number of construct(s) (e.g. depression, loneliness or happiness). The versatility of social media language data further allows for an unobtrusive and real-time assessment of wellbeing, based on the (longitudinal) type of data that are vastly available. The automatic analysis of such social media language data (SMTM) allows for wellbeing assessment at both individual and regional levels (e.g. by aggregating wellbeing scores across regions; Jaidka et al., 2020); therefore, both practitioners and policymakers can use SMTM method to inexpensively inform their treatments and interventions to increase human happiness/wellbeing.

Nonetheless, although such advantages of using SMTM exist, the development and the use of survey measures have already existed for a long period of time; therefore, the potential shortcomings and remedies to the problems that may arise are better known to the field. As also observed in our study's results, wording-related biases in surveys can be corrected with relatively easy to access and known methods (e.g. defining methods factor in a structural equation model). Further, the deployment of survey measures may require less expertise (e.g. pen–paper format), and the interpretation of the results can be easier (for instance, summing up item scores) while analysing language data may require more expertise and knowledge (e.g. pre-processing techniques, algorithm training methods).

Conclusion

The present study assessed the similarities and differences of wellbeing as assessed through both survey and social media language features. Both survey scores and social media language

features can be used to assess wellbeing, as shown by the theoretically relevant and largely similar wellbeing dimensions found in each network based on the two data types. Language and surveys also seem to provide slightly different information as reflected by the unique dimensions that were found in each network, which can allow for providing a more exhaustive way of measuring wellbeing if combined with each other.

CONFLICT OF INTEREST STATEMENT
None.

DATA AVAILABILITY STATEMENT
Data available on request due to privacy/ethical restrictions.

ETHICS STATEMENT
Institutional Review Board at the University of Pennsylvania (IRB No. 813866).

ORCID
Selim Sametoğlu https://orcid.org/0000-0002-5447-5410

ENDNOTE
[1] As a post hoc sensitivity analysis, we have estimated our final survey- and language-based networks with different gamma values, which are 0, 0.25 and 0.5. The comparisons revealed that glasso algorithm resulted with the same regularisation level (lambda value) for all gamma values we have tested and provided identical network structures and dimensions.

REFERENCES

Abbott, R. A., Ploubidis, G. B., Huppert, F. A., Kuh, D., Wadsworth, M. E., & Croudace, T. J. (2006). Psychometric evaluation and predictive validity of Ryff's psychological well-being items in a UK birth cohort sample of women. *Health and Quality of Life Outcomes*, *4*(1), 1–16.

Anglim, J., Horwood, S., Smillie, L. D., Marrero, R. J., & Wood, J. K. (2020). Predicting psychological and subjective well-being from personality: A meta-analysis. *Psychological Bulletin*, *146*(4), 279–323. https://doi.org/10.1037/bul0000226

Bartels, M., & Boomsma, D. I. (2009). Born to be happy? The etiology of subjective well-being. *Behavior Genetics*, *39*(6), 605–615.

Baselmans, B. M., & Bartels, M. (2018). A genetic perspective on the relationship between eudaimonic–and hedonic well-being. *Scientific Reports*, *8*(1), 1–10.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B: Methodological*, *57*(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research*, *3*(Jan), 993–1022.

Bollen, K. A., & Stine, R. A. (1992). Bootstrapping goodness-of-fit measures in structural equation models. *Sociological Methods & Research*, *21*(2), 205–229.

Borsboom, D., & Cramer, A. O. (2013). Network analysis: An integrative approach to the structure of psychopathology. *Annual Review of Clinical Psychology*, *9*(1), 91–121.

Borsboom, D., Deserno, M. K., Rhemtulla, M., Epskamp, S., Fried, E. I., McNally, R. J., Robinaugh, D. J., Perugini, M., Dalege, J., Costantini, G., Isvoranu, A., Wysocki, A. C., van Borkulo, C. D., van Bork, R., & Waldorp, L. J. (2021). Network analysis of multivariate data in psychological science. *Nature Reviews Methods Primers*, *1*(1), 1–18.

Bringmann, L. F., Elmer, T., Epskamp, S., Krause, R. W., Schoch, D., Wichers, M., Wigman, J. T., & Snippe, E. (2019). What do centrality measures measure in psychological networks? *Journal of Abnormal Psychology*, *128*(8), 892.

Burns, R. A., & Machin, M. A. (2009). Investigating the structural validity of Ryff's psychological well-being scales across two samples. *Social Indicators Research*, *93*(2), 359–375.

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

Chapman, B., & Guven, C. (2016). Revisiting the relationship between marriage and wellbeing: Does marriage quality matter? *Journal of Happiness Studies*, *17*, 533–551.

Chen, J., & Chen, Z. (2008). Extended Bayesian information criteria for model selection with large model spaces. *Biometrika*, *95*(3), 759–771.

Christensen, A. P., Gross, G. M., Golino, H., Silvia, P. J., & Kwapil, T. R. (2019). Exploratory graph analysis of the multidimensional schizotypy scale. *Schizophrenia Research*, *206*, 43–51.

Constantin, M. A. (2018). *Sample size recommendations for estimating cross-sectional network models*. Tilburg University.

Costantini, G., Richetin, J., Preti, E., Casini, E., Epskamp, S., & Perugini, M. (2019). Stability and variability of personality networks. A tutorial on recent developments in network psychometrics. *Personality and Individual Differences*, *136*, 68–78.

Deci, E. L., & Ryan, R. M. (2008). Hedonia, eudaimonia, and well-being: An introduction. *Journal of Happiness Studies*, *9*(1), 1–11.

Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The satisfaction with life scale. *Journal of Personality Assessment*, *49*(1), 71–75. https://doi.org/10.1207/s15327752jpa4901_13

Diener, E., 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*(2), 143–156.

Disabato, D. J., Goodman, F. R., Kashdan, T. B., Short, J. L., & Jarden, A. (2016). Different types of well-being? A cross-cultural examination of hedonic and eudaimonic well-being. *Psychological Assessment*, *28*(5), 471.

Dolan, P., Layard, R., & Metcalfe, R. (2011). Measuring subjective well-being for public policy.

Edwards, A. L. (1957). *The social desirability variable in personality assessment and research* (pp. viii, 108). Dryden Press.

Eichstaedt, J. C., Kern, M. L., Yaden, D. B., Schwartz, H. A., Giorgi, S., Park, G., Hagan, C. A., Tobolsky, V. A., Smith, L. K., Buffone, A., Iwry, J., Seligman, M. E. P., & Ungar, L. H. (2021). Closed- and open-vocabulary approaches to text analysis: A review, quantitative comparison, and recommendations. *Psychological Methods*, *26*(4), 398–427.

Eichstaedt, J. C., Smith, R. J., Merchant, R. M., Ungar, L. H., Crutchley, P., Preoţiuc-Pietro, D., Asch, D. A., & Schwartz, H. A. (2018). Facebook language predicts depression in medical records. *Proceedings of the National Academy of Sciences*, *115*(44), 11203–11208.

ElSherief, M., Kulkarni, V., Nguyen, D., Wang, W. Y., & Belding, E. (2018). Hate lingo: A target-based linguistic analysis of hate speech in social media. In *Proceedings of the international AAAI conference on web and social media* (Vol. 12, No. 1).

Epskamp, S., Borsboom, D., & Fried, E. I. (2018). Estimating psychological networks and their accuracy: A tutorial paper. *Behavior Research Methods*, *50*(1), 195–212. https://doi.org/10.3758/s13428-017-0862-1

Fernandes, H. M., Vasconcelos-Raposo, J., & Teixeira, C. M. (2010). Preliminary analysis of the psychometric properties of Ryff's scales of psychological well-being in Portuguese adolescents. *The Spanish Journal of Psychology*, *13*(2), 1032–1043.

Foygel, R., & Drton, M. (2010). Extended Bayesian information criteria for Gaussian graphical models. *Advances in Neural Information Processing Systems*, *23*.

Friedman, J., Hastie, T., & Tibshirani, R. (2008). Sparse inverse covariance estimation with the graphical lasso. *Biostatistics*, *9*(3), 432–441. https://doi.org/10.1093/biostatistics/kxm045

Garcia-Pardina, A., Abad, F. J., Christensen, A. P., Golino, H., & Garrido, L. E. (2022). Dimensionality assessment in the presence of wording effects: A network psychometric and factorial approach.

Giuntoli, L., & Vidotto, G. (2021). Exploring Diener's multidimensional conceptualization of well-being through network psychometrics. *Psychological Reports*, *124*(2), 896–919.

Golino, H., & Demetriou, A. (2017). Estimating the dimensionality of intelligence like data using exploratory graph analysis. *Intelligence*, *62*, 54–70.

Golino, H., & Epskamp, S. (2017). Exploratory graph analysis: A new approach for estimating the number of dimensions in psychological research. *PLoS ONE*, *12*(6), e0174035.

Golino, H., Shi, D., Christensen, A. P., Garrido, L. E., Nieto, M. D., Sadana, R., Thiyagarajan, J. A., & Martinez-Molina, A. (2020). Investigating the performance of exploratory graph analysis and traditional techniques to identify the number of latent factors: A simulation and tutorial. *Psychological Methods*, *25*(3), 292–320. https://doi.org/10.1037/met0000255

Guntuku, S. C., Yaden, D. B., Kern, M. L., Ungar, L. H., & Eichstaedt, J. C. (2017). Detecting depression and mental illness on social media: An integrative review. *Current Opinion in Behavioral Sciences*, *18*, 43–49.

Hargittai, E. (2020). Potential biases in big data: Omitted voices on social media. *Social Science Computer Review*, *38*(1), 10–24. https://doi.org/10.1177/0894439318788322

Hargittai, E., Füchslin, T., & Schäfer, M. S. (2018). How do young adults engage with science and research on social media? Some preliminary findings and an agenda for future research. *Social Media+ Society*, *4*(3), 2056305118797720. https://doi.org/10.1177/2056305118797720

Haslbeck, J. M. B., & Fried, E. I. (2017). How predictable are symptoms in psychopathological networks? A reanalysis of 18 published datasets. *Psychological Medicine*, *47*(16), 2767–2776. https://doi.org/10.1017/S0033291717001258

Henn, C. M., Hill, C., & Jorgensen, L. I. (2016). An investigation into the factor structure of the Ryff scales of psychological well-being. *SA Journal of Industrial Psychology*, *42*(1), 1–12.

Hsu, H. Y., Hsu, T. L., Lee, K., & Wolff, L. (2017). Evaluating the construct validity of Ryff's scales of psychological well-being using exploratory structural equation modeling. *Journal of Psychoeducational Assessment*, *35*(6), 633–638.

Hui, B. P. H., Ng, J. C. K., Berzaghi, E., Cunningham-Amos, L. A., & Kogan, A. (2020). Rewards of kindness? A meta-analysis of the link between prosociality and well-being. *Psychological Bulletin*, *146*(12), 1084–1116.

Humphries, M. D., & Gurney, K. (2008). Network 'small-world-ness': A quantitative method for determining canonical network equivalence. *PLoS ONE*, *3*(4), e0002051.

Isvoranu, A.-M., Epskamp, S., Waldorp, L., & Borsboom, D. (2022). *Network psychometrics with R: A guide for behavioral and social scientists*. Routledge.

Jaidka, K., Giorgi, S., Schwartz, H. A., Kern, M. L., Ungar, L. H., & Eichstaedt, J. C. (2020). Estimating geographic subjective well-being from Twitter: A comparison of dictionary and data-driven language methods. *Proceedings of the National Academy of Sciences*, *117*(19), 10165–10171.

James, P., Kim, E. S., Kubzansky, L. D., Zevon, E. S., Trudel-Fitzgerald, C., & Grodstein, F. (2019). Optimism and healthy aging in women. *American Journal of Preventive Medicine*, *56*(1), 116–124.

Jones, P. (2022). Networktools: Tools for identifying important nodes in networks (1.5.0). https://CRAN.R-project.org/package=networktools

Jones, P. J., Mair, P., & McNally, R. J. (2018). Visualizing psychological networks: A tutorial in R. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.01742

Kern, M. L., Benson, L., Steinberg, E. A., & Steinberg, L. (2016). The EPOCH measure of adolescent well-being. *Psychological Assessment*, *28*(5), 586.

Kern, M. L., Park, G., Eichstaedt, J. C., Schwartz, H. A., Sap, M., Smith, L. K., & Ungar, L. H. (2016). Gaining insights from social media language: Methodologies and challenges. *Psychological Methods*, *21*(4), 507–525.

Kern, M. L., Zeng, G., Hou, H., & Peng, K. (2019). The Chinese version of the EPOCH measure of adolescent well-being: Testing cross-cultural measurement invariance. *Journal of Psychoeducational Assessment*, *37*(6), 757–769.

Keyes, C. L., Myers, J. M., & Kendler, K. S. (2010). The structure of the genetic and environmental influences on mental well-being. *American Journal of Public Health*, *100*(12), 2379–2384.

Kim, E. S., James, P., Zevon, E. S., Trudel-Fitzgerald, C., Kubzansky, L. D., & Grodstein, F. (2019). Optimism and healthy aging in women and men. *American Journal of Epidemiology*, *188*(6), 1084–1091. https://doi.org/10.1093/aje/kwz056

Kjellstrom, S., & Golino, H. (2019). Mining concepts of health responsibility using text mining and exploratory graph analysis. *Scandinavian Journal of Occupational Therapy*, *26*(6), 395–410.

Longo, Y., Coyne, I., Joseph, S., & Gustavsson, P. (2016). Support for a general factor of well-being. *Personality and Individual Differences*, *100*, 68–72.

Lucas, R. E., & Brent Donnellan, M. (2012). Estimating the reliability of single-item life satisfaction measures: Results from four national panel studies. *Social Indicators Research*, *105*(3), 323–331. https://doi.org/10.1007/s11205-011-9783-z

Maccagnan, A., Wren-Lewis, S., Brown, H., & Taylor, T. (2019). Wellbeing and society: Towards quantification of the co-benefits of wellbeing. *Social Indicators Research*, *141*(1), 217–243. https://doi.org/10.1007/s11205-017-1826-7

Maydeu-Olivares, A., & Coffman, D. L. (2006). Random intercept item factor analysis. *Psychological Methods*, *11*(4), 344.

McCallum, A. K. (2002). MALLET: A machine learning for language toolkit. http://mallet.cs.umass.edu

Newman, M. (2005). Power laws, Pareto distributions and Zipf's law. *Contemporary Physics*, *46*(5), 323–351.

Okabe-Miyamoto, K., & Lyubomirsky, S. (in press). Happiness shapes and is shaped by social cognition and social connection. In D. Carlston, K. Johnson, & K. Hugenberg (Eds.), *The Oxford handbook of social cognition* (2nd ed.). Oxford University Press.

Oswald, A. J., Proto, E., & Sgroi, D. (2015). Happiness and productivity. *Journal of Labor Economics*, *33*(4), 789–822.

Park, G., Schwartz, H. A., Eichstaedt, J. C., Kern, M. L., Kosinski, M., Stillwell, D. J., Ungar, L. H., & Seligman, M. E. (2015). Automatic personality assessment through social media language. *Journal of Personality and Social Psychology*, *108*(6), 934.

Pons, P., & Latapy, M. (2006). Computing communities in large networks using random walks. *Journal of Graph Algorithms and Applications*.

Proctor, C., & Tweed, R. (2016). Measuring eudaimonic well-being. In *Handbook of eudaimonic well-being* (pp. 277–294). Springer.

Ryff, C. D. (1989). Happiness is everything, or is it? Explorations on the meaning of psychological well-being. *Journal of Personality and Social Psychology*, *57*(6), 1069.

Sametoglu, S., Pelt, D., Ungar, L. H., & Bartels, M. (2022). The value of social media language for the assessment of wellbeing: A systematic review and meta-analysis.

Santini, Z. I., Becher, H., Jørgensen, M. B., Davidsen, M., Nielsen, L., Hinrichsen, C., Madsen, K. R., Meilstrup, C., Koyanagi, A., Stewart-Brown, S., McDaid, D., & Koushede, V. (2021). Economics of mental well-being: A prospective study estimating associated health care costs and sickness benefit transfers in Denmark. *The European Journal of Health Economics*, *22*(7), 1053–1065. https://doi.org/10.1007/s10198-021-01305-0

Santini, Z. I., Nielsen, L., Hinrichsen, C., Nelausen, M. K., Meilstrup, C., Koyanagi, A., McDaid, D., Lyubomirsky, S., VanderWeele, T., & Koushede, V. (2021). Mental health economics: A prospective study on psychological flourishing and associations with healthcare costs and sickness benefit transfers in Denmark. *Mental Health & Prevention*, *24*, 200222.

Schuman, H., & Presser, S. (1996). *Questions and answers in attitude surveys: Experiments on question form, wording, and context*. Sage.

Sears, L. E., Shi, Y., Coberley, C. R., & Pope, J. E. (2013). Overall well-being as a predictor of health care, productivity, and retention outcomes in a large employer. *Population Health Management*, *16*(6), 397–405.

Shi, Y., Sears, L. E., Coberley, C. R., & Pope, J. E. (2013). Classification of individual well-being scores for the determination of adverse health and productivity outcomes in employee populations. *Population Health Management*, *16*(2), 90–98.

Shiffman, S., Hufford, M., Hickcox, M., Paty, J. A., Gnys, M., & Kassel, J. D. (1997). Remember that? A comparison of real-time versus retrospective recall of smoking lapses. *Journal of Consulting and Clinical Psychology*, *65*(2), 292.

Sirigatti, S., Stefanile, C., Giannetti, E., Iani, L., Penzo, I., & Mazzeschi, A. (2009). Assessment of factor structure of Ryff's psychological well-being scales in Italian adolescents. *Bollettino Di Psicologia Applicata*, *259*(56), 30–50.

Springer, K. W., & Hauser, R. M. (2006). An assessment of the construct validity of Ryff's scales of psychological well-being: Method, mode, and measurement effects. *Social Science Research*, *35*(4), 1080–1102.

Springer, K. W., Hauser, R. M., & Freese, J. (2006). Bad news indeed for Ryff's six-factor model of well-being. *Social Science Research*, *35*(4), 1120–1131.

Steptoe, A. (2019). Happiness and health. *Annual Review of Public Health*, *40*, 339–359.

Stillwell, D. J., & Kosinski, M. (2004). myPersonality project: Example of successful utilization of online social networks for large-scale social research. *American Psychologist*, *59*(2), 93–104.

Stochl, J., Soneson, E., Wagner, A. P., Khandaker, G. M., Goodyer, I., & Jones, P. B. (2019). Identifying key targets for interventions to improve psychological wellbeing: Replicable results from four UK cohorts. *Psychological Medicine*, *49*(14), 2389–2396.

Tankovska, H. (2020). Number of social network users worldwide from 2017 to 2025. Statista. https://www.statista.com/statistics/278414/number-of-worldwide-social-network-users/

Tennant, R., Hiller, L., Fishwick, R., Platt, S., Joseph, S., Weich, S., Parkinson, J., Secker, J., & Stewart-Brown, S. (2007). The Warwick-Edinburgh mental well-being scale (WEMWBS): Development and UK validation. *Health and Quality of Life Outcomes*, *5*(1), 1–13.

Thomas, P. A., Liu, H., & Umberson, D. (2017). Family relationships and well-being. *Innovation in Aging*, *1*(3), igx025.

Triado, C., Villar, F., Solé, C., & Celdrán, M. (2007). Construct validity of Ryff's scale of psychological well-being in Spanish older adults. *Psychological Reports*, *100*(3_suppl), 1151–1164.

Van de Weijer, M., Landvreugd, A., Pelt, D., & Bartels, M. (2021). Connecting the dots: Using a network approach to study the well-being spectrum.

VanderWeele, T. J., Hawkley, L. C., & Cacioppo, J. T. (2012). On the reciprocal association between loneliness and subjective well-being. *American Journal of Epidemiology*, *176*(9), 777–784.

Walsh, L. C., Boehm, J. K., & Lyubomirsky, S. (2018). Does happiness promote career success? Revisiting the evidence. *Journal of Career Assessment*, *26*(2), 199–219.

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature*, *393*(6684), 440–442.

Zeng, G., Peng, K., & Hu, C.-P. (2019). The network structure of adolescent well-being traits: Results from a large-scale Chinese sample. *Frontiers in Psychology*, *10*(2), 783.

Zessin, U., Dickhäuser, O., & Garbade, S. (2015). The relationship between self-compassion and well-being: A meta-analysis. *Applied Psychology. Health and Well-Being*, *7*(3), 340–364.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.