A SHORT INTRODUCTION TO SUBJECTIVE WELL-BEING: ITS MEASUREMENT, CORRELATES AND POLICY USES

Prepared for the international conference

Is happiness measurable and what do those measures mean for policy?

Organised by the Bank of Italy, the Centre for Economic & International Studies (CEIS), the Joint Research Centre of the European Commission and the Organisation for Economic Co-operation and Development (OECD)

2-3 April 2007, University of Rome ‘Tor Vergata’

André van Hoorn
Nijmegen Center for Economics (NiCE), Radboud University Nijmegen

Based on joint work with Stefano Castriota. Helpful comments by Anat Itay, Ramzi Mabsout and Robbert Maseland are gratefully acknowledged. The usual disclaimer applies.
1. **INTRODUCTION**

Few people have ever doubted that happiness is very important. In fact, starting at least with the Ancient Greeks, the concept has been subject of unremitting debate. Surely this would not have been the case if people generally felt it did not matter.¹

Since happiness has captured, and continues to capture, the interest of so many people, it should not come as a surprise that philosophers and many others debating the concept have long yearned for a way to measure it. The breakthrough came in the 1950s. Psychologists – until then mainly interested in negative emotional states such as depression and anxiety – became interested in positive emotions and feelings of well-being. Within the discipline a consensus grew that self-reports on how well life is going, can convey important information on underlying emotional states, and so the field pushed ahead with measuring what is best referred to as subjective well-being (commonly abbreviated as SWB). SWB, we should immediately note, is not the same as happiness although the terms are often used synonymously. SWB, in fact, is ‘a broad category of phenomena that includes people’s emotional responses, domain satisfactions, and global judgements of life satisfaction’ (Diener et al., 1999: p. 277). Specifically, reported SWB consists of two distinctive components (cf. Diener, 1994: p. 106): an affective part, which refers to both the presence of positive affect (PA) and the absence of negative affect (NA), and a cognitive part. The affective part is a hedonic evaluation guided by emotions and feelings, while the cognitive part is an information-based appraisal of one’s life for which people judge the extent to which their life so far measures up to their expectations and resembles their envisioned ‘ideal’ life.

Since the emergence of the field over five decades ago, the SWB literature has progressed rapidly. First, as recent surveys show, psychologists and other social scientists have taken huge steps in their understanding of the factors influencing people’s SWB. In addition, the methods by which empirical content is given to the concept of SWB have drastically improved and are expected to continue to do so as increasing use will be made of advances in information- and communication technology (ICT). As such, SWB research solicits increasing attention of politicians, government officials and the public alike.

This paper must be seen in light of the increasing attention for SWB and aims to provide a short introduction to the field. It starts with a discussion of the measurement of SWB in Section 2. In Section 3 we examine some factors underlying observed differences in SWB, both within and between nations. Section 4 explores some policy uses of SWB, notably in cost-benefit analyses and in the form of national SWB statistics. We conclude with some summarizing remarks in Section 5.

¹ King and Napa (1998) actually investigate how much value people attach to happiness by examining its contribution to the desirability of a certain hypothetical life and compare it with the contribution wealth and the presence of meaning make to the desirability of a given life. Their results confirm the folk wisdom developed over the ages: happiness and meaning make for the good life, and their effect on the desirability of a certain life is some five (happiness) to six (meaning) times higher than that of wealth. Relatedly, Diener and Oishi (2004) asked a sample of college students from different countries to rate the importance of happiness and other values on a scale from 1 to 9. Happiness came out first with a score of 8.0, slightly above health and love & affection (7.9), but well above wealth (6.8), amongst others. SWB further matters in that it fosters good outcomes in many domains of life, e.g. in work life, relationships and health (Lyubomirsky et al., 2005). Low SWB is also associated with general health risk and higher rates of suicide (Koivumaa-Honkanen et al., 2000 and 2001).
2. MEASUREMENT OF SUBJECTIVE WELL-BEING

2.1 Measurement scales and methods

The hallmark of measures of SWB is that they are obtained through self-reports: people are asked to evaluate their lives as a whole or some aspect of it. The questions can be relatively straightforward and a widely used one simply asks: ‘Taking all things together, would you say you are …: very happy, quite happy, not very happy or not at all happy’. More elaborate measures use multiple items to target a specific part of SWB and consequently render more reliable results single-item measures do (thought at an expense).

Much-used multi-item scales are the Positive and Negative Affect Schedule (PANAS) scales (Watson et al., 1988) and the Satisfaction With Life Scale (SWLS; Diener et al., 1985). The former provides a list of feelings and emotions and respondents are asked to indicate the extent to which they felt this way during a given period, e.g. the past week. The PANAS is very flexible in that it does not specify particular feelings and emotions and many different combinations are possible: interested, distressed, excited, upset, strong, guilty, et cetera. Items from SWLS solicit responses from individuals about the extent to which they agree or disagree with certain statements and it is designed specifically to capture satisfaction with life.\(^2\)

A critical aspect of measuring SWB is the method through which SWB reports are obtained. Two important measurement approaches are the Experience Sampling Method (ESM) or Ecological Momentary Assessment (EMA) (see Scollon et al., 2003 and Stone et al., 1999 for overviews) and the Day Reconstruction Method (DRM; Kahneman et al. 2004). The hallmark features of ESM/EMA are that they solicit frequent and immediate reports from individuals in their normal surroundings. These qualities, in turn, provide them with important advantages over other methods for soliciting self-appraisals (see Stone et al., 1999: pp. 27-28). Firstly, because individuals are in their natural settings when they complete the reports, the results are not distorted by unusual circumstances such as being in a laboratory; the method has high ‘ecological validity’. Furthermore, ESM and EMA do not have to be limited to self-reports but can also elicit randomly-timed reports on the details of people’s (momentary) environment. In subsequent analysis, environmental circumstances can therefore easily be related to the subjective emotional assessments. A second advantage is that having subjects complete the self-evaluation on the spot avoids retrospective distortion. A large literature shows that retrospective biases can be substantial and the use of momentary appraisals helps to limit them as much as possible. Finally, high-frequency assessment will increase the reliability and validity of measured SWB and improve its empirical analysis (see the discussion below).

ESM and EMA refer to a general method of measurement and do not specify a particular way of implementing it. Researchers are completely free in choosing the specific scale or scales used to measure SWB, e.g. a PANAS scale, a physiological test (for instance blood pressure) or a combination of different measures. In addition, the way in which the actual sampling takes place is not pre-specified, although practical difficulties of course set the boundaries on what is actually feasible. Modern information- and communication technologies offer interesting possibilities: when prompted to do so, respondents fill in the questionnaire send to their palmtops, and completed appraisals are included in a central database and analysed shortly after.

\(^2\) The statements are: (1) ‘In most ways my life is close to my ideal’, (2) ‘The conditions of my life are excellent’, (3) ‘I am satisfied with my life’, (4) ‘So far I have gotten the important things I want in life’, and (5) ‘If I could live my life over, I would change almost nothing’.
DRM is a more recently developed technique that shares most of the advantages of ESM/EMA but appears overall to be a more practical method. The technique asks subjects to keep a diary corresponding to events on or episodes of the day before. For the diary, they report on their experience, breaking it down in specific periods, e.g. by hour or by specific activities (commuting, household chores et cetera). In their research, Kahneman et al. (2004) simply ask respondents how they felt during each of the identified episodes, but again the method can be applied with any SWB scale.

2.2 Reliability and validity

Reliability
The reliability of an indicator can be defined as its overall quality, i.e. its consistency and its ability to give the same results in repeated measurement. The most outstanding feature of reliability is the test-retest correlation of the specific measure under scrutiny. It is well-known that minor differences in circumstances and technical features of the specific questionnaire used, affect the reported level of SWB (see Schwarz and Strack, 1999: p. 62). Correspondingly, the test-retest correlation for most single-item measures is about 0.40, reaching 0.60 when the same question is asked twice during the same one-hour interview (ibidem). In general, the reliability of SWB measures as evidenced by test-retest correlation is substantially lower than that found for common microeconomic variables such as personal income. It is important to realise, however, that test-retest correlation is not only substantially influenced by the lag between times of asking (e.g. one hour, two weeks), but also, and more importantly, by the specific measurement scale used. Not surprisingly, the more advanced measures such as multi-item questionnaires produce more reliable SWB scores.

Validity
Validity, not unlike reliability, comes in many sizes and shapes. In its regular use, validity denotes so-called ‘construct validity’ and refers to the actual meaning of a term or construct, in this case SWB (Larsen and Fredrickson, 1999: p. 44). That is, a measure has construct validity if it indeed is able to capture the construct, in this case SWB, it is intended to capture. Other important dimensions of validity, often contributing to an indicator’s overall construct validity, are convergent validity and discriminant validity. The former refers to the correspondence between the indicator of interest and other indicators purporting to measure the same construct. If the measure of interest is to have convergent validity the results obtained should converge on the results obtained using these other measures. Discriminant validity is somewhat the converse of convergent validity: a measure does well on this account if it is unrelated to measures it should not be related with, i.e. indicators purporting to measure some opposing construct.

The SWB literature pays a lot of attention to the validity of its measures. The conclusion in most of the reviews is that measures of SWB do quite well on the various dimensions of validity mentioned above (see, for instance, Di Tella and MacCulloch, 2006, Diener, 1994, Diener et al., 1999, Frey and Stutzer, 2002, Kahneman and Krueger, 2006, Nettle, 2005 and Layard, 2005). This is not to say, however, that SWB does not lack a universal definition. Indeed, as a concept, SWB as of yet does not have a specific theoretical basis and, accordingly, the field appears largely driven by empirical work: SWB is generally conceptualised through the specific indicators used, i.e. by the specific questions asked.

The most well-known challenge to measures of SWB and their validity concerns their demonstrated sensitivity to minor life events. Most importantly, whimsical circumstances – a famous example is the weather, whether it is rainy or sunny – have a significant impact on the
level of SWB people report (Schwarz and Strack, 1999: p. 62). Such effects are obviously an important weakness of SWB measures and, in fact, survey data in general (cf. Bertrand and Mullainathan, 2001). The errors generally seem to be of a random rather than a structural nature, however, and consequently do not mean the measurement instrument is systematically biased. Indeed, using large enough samples would go a long way in addressing possible problems introduced by contextual factors influencing the reported level of SWB. This is actually the advantage of using ESM, EMA or DRM; in repeated measures of SWB for a single individual, the influence of trivial events is lessened.

At the same time, SWB indicators do face a more substantial challenge to their validity, specifically with respect to their validity in cross-national comparisons. This challenge concerns the possible impact of culture and language on SWB ratings (see Wierzbicka, 2004). By the very nature of SWB and the way in which it is measured, i.e. through questionnaires, cultural and linguistic factors will introduce biases in country-level SWB scores. The extent to which they indeed do so is largely unexplored: overall, culture is found to matter (e.g. Vittersø et al., 2005) though it is not necessarily an explanatory factor in observed differences in SWB (e.g. Diener et al., 1995).

In addition, increasingly attention is paid to the relation between SWB on the one hand and human physiology and brain activity on the other. This work shows that variations in SWB are associated with concrete objectively observable biological phenomena. Ultimately, work along these lines may lead to an objective anchor for SWB that would improve the comparability of SWB levels across national and linguistic barriers. Obviously many issues – not least the question of causality – need resolving, before a reliable biological basis is given to observed variation in SWB, bu a very preliminary step in this direction has already been taken. In a recent study Blanchflower and Oswald (2007) relate differences in SWB across countries to differences in blood pressure.

3 CORRELATES AND DETERMINANTS OF SUBJECTIVE WELL-BEING

Surveys show that social sciences have taken huge steps in their understanding of the factors underlying differences in SWB ratings (e.g. Diener et al., 1999). This section briefly

---

3 Boulding (1972: p. 466) describes this effect of the law of large numbers as follows: ‘Even if we cannot get very reliable measures for the individual, by the famous principle which I have sometimes called “Katona’s Law”, that the summation of ignorance produces knowledge, we may find an operation or instrument with self-cancelling random factors which will give us a much better measure for a hundred or a thousand individuals than we can get for one.’

4 This view on cultural and linguistic factors as a source of bias in cross-country comparisons, is not meant to deny that the role of these factors in reported SWB itself is interesting and worthy of attention.

5 For example, using positron emission tomography (PET scans) of regional cerebral blood flow, Lane et al. (1997) found, amongst others, that pleasant and unpleasant emotions differed from neutral ones by statistically highly significant elevated blood flow in the vicinity of the medial prefrontal cortex, thalamus, hypothalamus and midbrain. Levesque et al. (2003) applied functional magnetic resonance imaging (fMRI) and found that sad feelings were associated with significant bilateral activations of the midbrain, the medial prefrontal cortex, the anterior temporal pole and the right ventrolateral prefrontal. Ryff et al. (2004) present some preliminary findings showing that people with higher SWB (specifically more meaning, purposeful engagement et cetera in their lives) have lower levels of daily salivary cortisol and pro-inflammatory cytokines. In addition, for them the duration of REM sleep is longer than for those with lower levels of reported SWB. Finally, Steptoe et al. (2005) find that positive affect is associated with reduced neuroendocrine, inflammatory and cardiovascular activity. Positive affect was also inversely related to cortisol output during the day (controlling for other factors such as age and gender) and heart rate. During mental stress testing in the laboratory people with higher positive affect had smaller plasma fibrinogen stress responses.

6 This paragraph draws heavily on Castriota (forthcoming).

7 See also Diener et al. (1995), Diener and Seligman (2004), Frey and Stutzer (2002), Lyubomirsky et al. (2005), Nettle (2005) and Layard (2005), amongst others. Earlier accounts of the state of affairs in SWB research are provided by Fellows (1966), Veenhoven (1984) and Wilson (1967).
discusses some of the uncovered correlates and determinants of SWB, classifying them in six broad groups: (i) personality factors; (ii) contextual and situational factors; (iii) demographic factors; (iv) institutional factors; (v) environmental factors; and (vi) economic factors.8

Psychologists have deeply studied the influence of personality on SWB, and found it to be the strongest and most dependable factor underlying differences in SWB between persons. In a famous study Tellegen et al. (1988) compared levels of SWB for monozygotic and dizygotic twins raised together and raised apart. Their study shows that 40% of the variance in positive emotionality and 55% of the variance in negative emotionality is attributable to genes, whereas shared familial circumstances account for only 22% and 2% of observed variance respectively. Much work has assessed the role of measured personality characteristics and these are also consistently found to be highly significant predictors of SWB. Notably neuroticism and extraversion go along way in accounting for differences in levels of SWB (see, e.g., Hayes and Joseph, 2003).

Although inherent factors play a fundamental role in SWB, individual, contextual and situational factors are also important sources of difference in SWB scores. Notably, a consistent finding across samples of individuals reporting on, amongst others, SWB is that better health is associated with higher SWB, and that married people report higher SWB than others, e.g. single people, divorced people et cetera.

The third group of factors strongly associated with SWB concerns demography. Gender and age in particular are robust determinants of SWB across samples. Generally, women report higher SWB scores than men do, and SWB is U-shaped with age: SWB is higher among young people, declines in middle age cohorts and increases again at older age.

Institutional conditions constitute a fourth group of factors found to have a systematic relationship with SWB. For instance, the results of Frey and Stutzer (2000) suggest that forms of direct democracy (e.g. referenda) increase the level of SWB. At a more abstract level, Radcliff (2001) finds a positive relation between the ideological complexion of governments and levels of SWB. He also reports a positive correlation between qualitative features of the welfare state and SWB. Finally, Veenhoven (2000) finds that political and private freedom add to SWB but only in rich countries.

Environmental conditions are an important factor in observed differences in SWB that operates strictly at a macro level. For example, Rehdanz and Maddison (2005), using data on 67 countries between 1972 and 2000, find that climate variables have a highly significant effect on SWB and that climate changes due to global warming might reduce SWB around the world in the next decades (see also the next section). In much the same fashion, the analysis by Becchetti et al. (2007) confirms the link between climate and SWB but indicates that global warming might, in contrast, lead to higher SWB worldwide.

Finally, part of individual and cross-country differences in well-being is attributable to differences in economic circumstances. The literature has developed a quite clear understanding of the role factors like unemployment and inflation play in SWB (e.g. Clark and Oswald, 1994; Winkellmann and Winkellmann, 1998; Di Tella et al., 2001; Becchetti et al., 2006). In particular, it is well-established that unemployment affects SWB through two channels: it has a direct negative effect on people who lose their job (keeping income constant), and an indirect negative effect on the entire population (higher risk of losing a job). The most significant of all economic variables seems to elude us, however. It is still not entirely clear what the role of GDP (gross domestic product) and GDP growth is. In this

---

8 This overview is necessarily brief. Accordingly, important factors like religious beliefs are not discussed, nor does the section pay attention to the key issue of causality - does high SWB lead to marriage or does marriage lead to high SWB.
context Abramovitz (1979: p. 7) discusses the findings of Easterlin (1974) and first mentions the ‘Easterlin paradox’. According to this paradox, within a country richer people on average report higher SWB than poorer people in the same country do, whereas a comparison between countries reveals only a minor relation between income levels and SWB. Moreover, within nations an increase in per-capita income over time is not associated with higher SWB. Much effort has been put in further investigation of Easterlin’s paradox. This work has robustly established the importance of relative income, but for aggregate-level income no clear picture emerges. A particularly interesting finding is by Di Tella et al. (2003). They examine time-series data showing that GDP and GDP growth are significantly correlated to SWB, but only when down-trending time dummies are included in the regression equation.9

4. POLICY USES OF SUBJECTIVE WELL-BEING

Following the major advances in the field of SWB, much work is currently underway that evaluates the contribution it can make to shaping and appraising policy. More concretely, several ideas exist for the construction of ‘a national index of subjective well-being’ or a related indicator, which would subsequently serve as a key policy goal. In addition, people are experimenting with the use of SWB data to assess the costs and benefits of policy alternatives. We discuss both policy uses next.10

4.1 Towards a comprehensive picture of well-being

Frameworks designed to give a comprehensive account of how well a nation and its citizens are doing generally discern three key areas. Measures of economic performance are complemented with indicators on important social and environmental issues. For example, the Belgian Federal Planning Bureau (2005) in its insightful TransGovern model regards society’s access to human, environmental and economic resources as determining the evolution of well-being. Correspondingly, it discerns three basic capitals of development:

1. **Human capital:** comprising the standard of living (material well-being), health (both mental and physical) and knowledge/capacities (what individuals know and are able to do);
2. **Environmental capital:** including both natural resources (water, air, land and mineral resources) and the biosphere with all its biological diversity;
3. **Economic capital:** subdivided in physical and technological capital (equipments, buildings, infrastructure, and intangible assets including software and technology patents) and net financial assets.

Typically, well-being frameworks such as this one include only objective indicators; whether it be GDP or related measures of economic activity, water quality or telephone connections, illiteracy rates or patent applications, suicide numbers or victimisation rates, traffic accidents or life expectancy, there is no doubt that from these indicators much can be learned about the quality of life. At the same time, there is growing recognition that, though insightful, there is more to well-being than these measures are able to capture.

---

9 For further evidence on the importance of both relative and absolute income see, amongst others, Clark and Oswald (1996), Diener et al. (1995), Easterlin (1995) and McBride (2001).
10 Other uses not discussed here are, amongst others, as an alternative to the behaviourist approach to analysing the effects of taxes (e.g. Gruber and Mullainathan, 2005) and in the construction of equivalence scales (e.g. van Praag van der Sar, 1988). Further note that, also not elaborated on here, interpersonal comparability is implicit in virtually all policy use of indicators, whether objective or subjective.
The central tenet of SWB research is that the study of well-being cannot get around measuring people’s cognitive and affective reactions to life as whole or specific domains thereof (Diener and Suh, 1997: p. 200). Most importantly, SWB is able to capture’s people actual experience in a direct manner, while economic, social and environmental indicators do so only indirectly (p. 205). This, in turn, matters because what is experienced does not have to coincide with objective conditions, and indeed a large deviation may be observed. In fact, it is often argued SWB indicators are useful complements to objective indicators precisely because there is a divergence between what people (reportedly) experience on the one hand and with what is captured in the objective indicators on the other (see, in particular, Diener and Seligman, 2004: pp. 2-3).

Building on the above-discussed methods for measuring SWB and the perceived need that something important is missing in the current set of well-being indicators, researchers in the field have given the idea of national SWB statistics serious consideration (Diener, 2000 and 2006; Kahneman et al., 2004). To give a taste, a nationally representative sample could be given palmtops and asked to report on their SWB using ESM/EMA or DRM. The cross-section would comprise different age groups, geographical regions, occupational categories and income levels Diener (2000: p. 40). The actual survey items could further target all components of SWB, the affective component (the presence of PA and the absence of NA) and the cognitive component so as to maximise the usefulness, reliability and validity of SWB thus measured. A ‘national index of subjective well-being’ or ‘national well-being accounts’ constructed along these lines promises to be a very useful complement to existing objective indicators and to add greatly to a more comprehensive picture of the quality of life.

4.2 The measuring-rod of happiness

Pigou (1952: p. 11) famously thought of money as, amongst others, a ‘measuring-rod’, defining economic welfare as ‘that part of social welfare that can be brought directly or indirectly into relation with the measuring-rod of money’. For much of applied economics, money indeed is the measuring-rod Pigou foisted upon it. In many policy areas, public goods and externalities being the ones most familiar to economists, there is just no way of putting a price tag on the relevant dimensions of the different options available, however. In some of these cases subsequently, SWB can provide a way to deal with the problem of missing prices. Below we discuss three examples of how SWB can be used to put a price on certain developments or externalities otherwise escaping (monetary) valuation.12

Van Praag and Baarsma (2005) address a concrete valuation problem. They combine a more mainstream approach to valuing an intangible, viz. noise nuisance from an airport, with the use of SWB data. Ordinarily, a hedonic regression analysis of house prices is used to calculate the shadow costs of noise pollution. The hedonic approach to valuing intangibles is prone to leave some residual costs, however, and these can be valued using measured SWB. The method thus developed is particularly useful when for some reason, notably market imperfections, price differences do not fully capture the costs of an externality.

An important area in which SWB data can usefully be applied to circumvent the problem of missing prices is in environmental valuation issues. Notably, Welsch (2006) uses panel data on ten European countries to analyse the effect of air pollution on SWB. Controlling for income, he finds that differences in SWB, both between countries and within

---

11 Supporting the field of SWB’s approach to well-being is Sumner’s (1995) argument that an adequate theory of the nature of welfare or well-being cannot be objective and hence must be subjective.

12 More well-known, contingent valuation provides a further approach to solving the problem of missing prices (see the symposium in the Fall 1994 edition of the Journal of Economic Perspectives for a thorough introduction).
countries over time, can partially be attributed to environmental quality. The improvement in air quality over the period 1990 – 1997 is valued at roughly $750 per capita per year in the case of nitrogen dioxide and about $1400 per capita per year in the case of lead (averaged over all countries).

Finally, Frey et al. (2007) notably expand the dimensions of life to which valuation techniques have been applied and assess the costs of terrorism using SWB data. The case applies to the Republic of Ireland, during the period 1970-1999. Data is taken from the Eurobarometer survey, and in the analysis differences in life satisfaction are explained from levels of terrorism, household income and other personal characteristics. Their findings reveal that people, on average, would be willing to give up some 41% of their income to have the level of terrorism reduced to the level prevailing in more peaceful parts of the country.

5. SUMMARISING REMARKS AND CONCLUSION

This paper has presented a very brief introduction to the field of SWB or subjective well-being along three lines, its measurement, its correlates and its policy uses. Key points, corresponding to the three aspects of SWB research discussed, are:

1. SWB measures are meaningful in the sense that they are indeed able to provide valid and reliable information on how well people and societies as a whole are doing;
2. a broad range of factors are found to correlate with or be a causal factor in SWB, both at the level of individuals and that of countries;
3. SWB data can be used to shape and appraise policy.

These conclusions notwithstanding, it must be pointed out that important challenges remain. Notably, work on SWB thus far has not been very systematic and in important areas the current understanding is limited. The implication is that much work remains to be done before SWB is able to play the role in policy many authors envision for it.13

However, as SWB research is a rapidly developing field, many weaknesses of the subject area are likely to be addressed in the near future. For instance, linking SWB measures to objective, biological phenomena could ultimately help cross-country comparison of SWB scores, which, in turn, could inform public policy (although, of course, without providing a recipe for government action). Similarly, some areas have already been identified in which measures of SWB can complement existing measures and extent the scope of cost-benefit analyses. With more such work underway, SWB indeed appears a very promising area of research, from the perspective of politicians, government officials and the public alike.

13 And even then it is hard to think of SWB as being the only relevant or all-encompassing measure.
REFERENCES


