The Measurement of Subjective Well-Being in Survey Research
Vol. 16, No. 3, pp. 335–372

CONCEPTUALIZING SUBJECTIVE WELL-BEING AND ITS MANY DIMENSIONS – IMPLICATIONS FOR DATA COLLECTION IN OFFICIAL STATISTICS AND FOR POLICY RELEVANCE

Christopher Mackie[1], **Conal Smith**[2]

ABSTRACT

Subjective well-being encompasses several distinct but interacting aspects of people's feelings, attitudes, and experiences. This paper assesses the state of the art for measuring these dimensions of people's lives, which typically involves analyzing self-reports of subjective well-being collected in survey instruments; however, other potentially complementary, technology-driven tools are emerging as well. We first answer the question, "what is subjective well-being?" and unpack its multidimensionality. The role of national statistics offices in measuring subjective well-being and deriving official statistics is considered next. We conclude by discussing how different characteristics of well-being constructs shape their applicability to policy. The overarching conclusion is that–while methodological limitations are present and a number of fundamental research challenges remain–understanding of how to collect and interpret data on subjective well-being has made enormous strides in the last two decades, and policies for a wide range of domains are beginning to be usefully informed.

Key words: subjective well-being, national statistics, policy.

1. Introduction

Notions of subjective well-being (SWB) or happiness have a long tradition as central elements of the good life. However, until recently, these concepts were often deemed impossible to measure, and certainly beyond the scope of official statistics. In the past two decades, however, an increasing body of evidence has shown that SWB can be measured in surveys, that such measures are valid and reliable, and that they can inform policy making. This evidence has been reflected in the exponential growth of research in this field. As documented by Krueger and

[1] Committee on National Statistics, National Academies – Washington, DC. E-mail: cmackie@nas.edu.
[2] Organisation for Economic Co-operation and Development – Paris. E-mail: Conal.SMITH@oecd.org.

Schkade (2008), OECD (2013), and elsewhere, a near exponential growth in the publication of articles on happiness and related subjects has emerged since the 1990s, including in the top economics journals.[3]

Reflecting increasing interest in SWB from researchers, policy-makers and the public—and further adding to its legitimacy—the report of the *Commission on the Measurement of Economic Performance and Social Progress* (Stiglitz et al., 2009) recommended that national statistical agencies collect and publish measures of SWB. This was followed in 2013 by the publication of the *OECD Guidelines on the Measurement of Subjective Well-being*, aimed at encouraging the collection and publication of such measures by national statistical offices. In the same year the National Academy of Sciences published *Measuring Happiness, Suffering, and Other Dimensions of Experience*, which investigated the application of experienced well-being measures to policy in the United States.

A large number of national statistical offices are now collecting SWB measures either on an experimental basis or as part of their core programs. Among OECD countries, 32 out of 34 national statistical offices collect data on life satisfaction along the lines recommended by the OECD *Guidelines*. The United Kingdom, for example, now collects four measures of SWB aligned with the OECD *Guidelines* in its *Annual Population Survey*, providing a total sample of approximately 160,000 each year. On 3 September 2014, the UK Statistics Authority granted these four measures accredited National Statistics status, confirming them as part of the highest tier of official statistics in the UK.

With the increasing prominence of SWB in official statistics, it is useful to review their conceptual scope and to consider how they can be applied to policy. While the majority of measurement initiatives and academic research have focused on how people evaluate their lives (often, misleadingly, described as measures of "happiness"), widespread consensus has emerged that SWB has multiple distinct dimensions. This paper presents a general overview of SWB and its underlying complexity, then discusses the implications for data collection, measurement, and informing policy.

2. What is subjective well-being?

Subjective well-being encompasses several separate but interacting aspects of people's feelings, attitudes, and experiences. The construct covers a number of different aspects of a person's subjective state; however, there is debate about exactly what elements should be emphasized (Diener et al., 1999; Kahneman,

[3] The Journal of Economic Literature, the Journal of Economic Perspectives and the Journal of Political Economy have all published papers on subjective well-being in the recent past (Frey and Stutzer, 2002; Di Tella, MacCulloch, and Oswald, 2001; Rayo and Becker, 2007, among many others).

Diener, and Schwarz, 1999). Kahneman and Krueger (2006), for example, focus primarily on experienced well-being, while Huppert et al. (2009) emphasize measures of good psychological functioning. The OECD *Guidelines* (2013) defines SWB as involving "good mental states, including all of the various evaluations, positive and negative, that people make of their lives, and the affective reactions of people to their experiences[4]." This characterization is inclusive in nature, encompassing a broad dimensional conceptualization of SWB. In particular, the reference to good mental functioning acknowledges concepts such as interest, engagement, and meaning alongside more commonly identified notions of satisfaction and emotional state. Similarly, Diener (2006) argues that "subjective well-being is an umbrella term for the different valuations people make regarding their lives, the events happening to them, their bodies and minds, and the circumstances in which they live."

Two definitional points are worth making here. First, SWB is narrower in scope than are self-reported measures in general which may be directed toward outcomes that have no relationship to mental states. For example, a survey may ask respondents to report income, marital status, or employment information, none of which is directly a measure of SWB (though they certainly may be correlates to it). Second, SWB is not necessarily synonymous with well-being as a whole. In the measurement of human welfare, largely non-subjective variables such as income levels, health status, knowledge and skills, environmental quality and social connections often play important roles.

In order to develop meaningful measures of SWB, it is essential to identify which of its elements is to be the central focus. Although some researchers argue in favour of a single overall construct (Stewart-Brown and Janmohamed, 2008), SWB is more commonly acknowledged to encompasses three core dimensions—life evaluation, experienced or hedonic well-being, and eudaimonia (which includes concepts such as sense of purpose or meaning and locus of control)—and that each should be measured. Although both the OECD *Guidelines* (2013) and NAS (2013) recommend measuring these dimensions separately, they also recognize that they are interrelated. Many of the distinctions in SWB measurement constructs relate to their temporal characteristics which may be thought of in terms of a continuum, with essentially real-time assessments of experience, emotional state, or sensations at the shortest end of the spectrum and overall evaluations of life satisfaction, purpose, or suffering at the other end (the longest reference period). Sense of meaning or purpose may impact a respondent's assessment of either a momentary situation (why do I not mind reading Dr. Seuss to my child over and over again?) or to life evaluation (will studying 15 hours a day to become a physician lead to a better life?) (NAS, 2013, p. 15). Next, we define and review the major dimensions of SWB—life satisfaction, affect and experienced well-being, and eudaimonia.

[4] This definition derives from Diener et al. (2006).

2.1. Life evaluation

Life evaluation is conceived of as resulting from a reflective assessment of a person's life. Such assessments entail judgement by an individual, and stand in contrast to descriptions of mood or emotional state. Pavot et al. (1991) describe the evaluative process as involving individuals constructing a "standard" that they perceive as appropriate for themselves, and then comparing the circumstances of their life to that standard. Although it is not clear whether this process of comparison is a conscious one, in practice, the relatively short response time associated with life evaluation questions in surveys suggests that respondents will typically use a heuristic to form a rating (OECD, 2013).

It is tempting to equate life evaluation with an economist's definition of utility as the criteria by which different choices are evaluated. There is a prima facie plausibility to the idea that people pursue goals that maximise the evaluation of their lives, a view that has found significant empirical support (Clark, 2001; Clark, Frijters, and Shields, 2008; Frijters, 2000; Helliwell and Barrington-Leigh, 2010). However, there are also strong reasons to be cautious in treating measures of life evaluation as measures of utility. First, although economists traditionally assume (at least implicitly) that the remembered utility on which people base their decisions is equivalent to the sum of momentary utilities associated with moment-by-moment experience, SWB measurement has revealed this to not always be the case. Life evaluations are based in part on how people remember their experiences, which can differ significantly from how they actually experienced things at the time (Kahneman et al., 1999). For example, the so-called "peak-end rule" states that a person's evaluation of an event is based disproportionately on the most intense (peak) and last (end) emotions experienced during the event, rather than the average or integral of emotional experiences over time. A second critique of the view that life evaluation measures utility focuses on the observation that people are prepared to trade off life satisfaction in order to achieve other outcomes. If measures of life evaluation fully captured utility as conceived of by economists, the notion of people accepting reduced levels of life evaluation in order to gain some other goal would make little sense.[5] Despite these concerns, measures of life evaluation remain of high interest for two reasons. First, although life evaluation is probably not measuring an economist's conception of utility, other approaches to analysing utility also have limitations. Measures of life evaluation can therefore add to the sum total of knowledge without themselves being perfect measures of utility. Second, regardless of whether life evaluations measure utility, how people feel about their lives is an important consideration in its own right. Life evaluations may provide insights into people's well-being more generally, even if they do not align perfectly with some over-arching view of the concept.

[5] Benjamin, Kimball, Heffetz, and Rees-Jones (2013) examine such trade-offs in the context of stated preferences in the residency choices of medical students.

As a reflective construct of a respondent's subjective state, life evaluation is usually measured through one or more survey questions. Perhaps the best validated measure in the psychological literature is the 5 question Satisfaction With Life Scale developed by Diener and Pavot (1993). However, the two most widely used measures are single item survey questions: the Self Anchoring Striving Scale (more commonly known as the Cantril Ladder) and the *World Values Survey* Satisfaction With Life question. The Cantril Ladder is used in the *Gallup World Poll* and is thus the basis of much recent research on the drivers of life evaluation across countries. Until recently it was believed that the Cantril Ladder and the Satisfaction With Life question collected slightly different information with the former being the more purely evaluative of the two. Recently, however, evidence from the Gallup World Poll and research based on split sample surveys in the UK *Household Opinion Survey* has provided convincing evidence that the two questions are closely comparable (Helliwell, Layard, and Sachs, 2013; ONS, 2011). In part due to this finding, OECD (2013) recommends a life evaluation question based on the simpler World Values Survey version:

> *The following question asks how satisfied you feel on a scale from 0 to 10. Zero means you feel "not at all satisfied" and 10 means you feel "completely satisfied".*
>
> *Overall, how satisfied are you with life as a whole these days?" [0-10]*

This question, with minor variations, forms the basis of most subjective well-being data currently collected by national statistical agencies.

2.2. Affect, experienced well-being

Affect is the term psychologists use to describe a person's feelings. Affect can be thought of as particular feelings or emotional states and is typically measured with reference to a particular point in time. Such measures capture how people experience life rather than how they remember it (Kahneman and Krueger, 2006). While an overall evaluation of life can be captured in a single measure, affect has at least two distinct−so-called hedonic−dimensions associated with positives and negatives (Kahneman et al., 1999; Diener et al., 1999). Positive affect captures emotions such as happiness, joy, and contentment. Negative affect comprises the experience of unpleasant emotional states such as sadness, anger, fear, and anxiety. While positive affect is thought to be largely uni-dimensional (in the sense that positive emotions are strongly correlated with each other), negative affect is more multi-dimensional. For example, it is possible to imagine at a given moment feeling anger but not fear or sadness.

Bradburn (1969) was one of the first researchers to determine that positive and negative affect are not opposite ends of one dimension but are largely independent of one another; a person can rate highly on one state and either high or low on the other. Bradburn's findings have been replicated many times; for

example, Gere and Schimmack (2011) found that, even after correcting for measurement error and bias, positive and negative feelings were distinct. This body of research evidence led the NAS panel to conclude that:

> Both positive and negative emotions must be accounted for in experienced well-being measurement, as research shows that they do not simply move in an inverse way. For example, an activity may produce both negative and positive feelings in a person, or certain individuals may be predisposed to experience both positives and negatives more strongly. Therefore, assessments of [experienced well-being] should include both positive and negative dimensions in order for meaningful inferences to be drawn (p. 39).

Other dimensions of experienced well-being such as arousal, which relate to positive and negative emotions in a range of ways, are important as well. Sensations such as pain, numbness, heat, or cold may also figure into emotional states and into hedonic assessment of those states−particularly if the context is people's health or housing conditions. Certainly, people experiencing pain will on average report higher levels of negative well-being, all else being equal (Krueger and Stone, 2008).

The term hedonic well-being typically is used in association with the emotional (or affect) component SWB. And, although the term "experienced well-being" is sometimes treated synonymously with the affect, they are not identical. Experienced well-being is broader in the sense that it may include pain and other sensations that factor into suffering or happiness which may be omitted by the narrower hedonic focus on emotions. Even more broadly, as described below, appraisals of concepts beyond the emotional, such as meaning or purpose, may also be included in the experienced well-being construct (NAS, 2013). Measuring "experience" broadly is essential for addressing issues of long-term suffering which are of concern to policy makers. As elaborated below, these characteristics carry also implications for data collection strategies.

The characteristics of affective states also raise an interesting question about their relationship to life evaluation. Research has established that positive and negative experience track at least partially independently of life satisfaction and of each other. Kahneman et al., (1999) argue for the existence of a "good/bad" axis on which people are able to place experiences based on their emotional states at the time. In principle, this process is similar to that involved in forming life evaluations from remembered affective states. Kahneman's point is that affective states can be compared and that one can therefore reasonably aggregate measures of current affect. For this reason, affect measures are sometimes reported in terms of affect balance, which captures the net balance between positive and negative affect (Kahneman and Krueger, 2006).

A number of measurement approaches have been used to measure affect and, more broadly, experienced well-being. Sometimes approach is dictated by the

measurement objective; sometimes it is constrained by survey (or other) data gathering instrument. The basic categories are:

- Ecological momentary assessment, which signals a person to respond in the moment. Sometimes considered the gold standard for measuring affect, in the experience sampling method (ESM), participants are prompted to record their feelings and perhaps the activity they are undertaking at either random or fixed time points, usually several times a day, throughout the study period, which can last several weeks. To maximise response rates and ensure compliance throughout the day, electronic diaries are often used to record the time of response. While the ESM produces an accurate record of affect, it is also expensive to implement and intrusive for respondents.
- Reconstructed activity-based measures; time use/day reconstruction methods (DRM) allow contextual information to be linked to measures associated with specific activities (e.g. job search, child care, commuting) and in turn to policy questions. DRM, in which respondents are questioned about events from a time-use diary recorded on the previous day, are often more practical and viable for government surveys. Research has shown that the DRM produces results comparable with ESM, but with a much lower respondent burden (Kahneman et al., 2004).
- Single day measures, which ask respondents about their experiences globally for a given day or episodes during that day. Surveys are typically administered at the end of day or the next day. A number of important survey and measurement issues arise when single day measures are used to approximate results of momentary measures (these are dealt with in the paper by Dylan Smith in this volume).

Experience Sampling, DRM, and similar methods for collecting affect data in time-use studies allow for analyses capable of associating particular affective states with specific activities. Measures of affect collected in this way thus capture well-being as reported by a person in a particular place, time, and set of circumstances as opposed to some sense of overall SWB. It is also possible to collect affect data in general household surveys via questions about a person's mood or emotional state over a particular recall period.[6] However, although such measures capture information on a person's affective state, they cannot easily capture information linking affect to particular activities. On the other hand, it is also possible to collect information about some aspects of eudaimonia (see the next section) using similar techniques to those used to measure experienced affect. For example, the American Time Use Survey well-being module collects information on experienced "meaning and purpose" associated with specific daily episodes.

[6] The Gallup World Poll contains a range of questions on affect during the previous day, which have been extensively tested. The UK Office of National Statistics has collected similar measures of affect in its Integrated Household Survey programme.

2.3. Eudaimonia

A substantial literature exists on the concept of good psychological functioning, sometimes referred to as "flourishing" or "eudaimonic" well-being (Huppert et al., 2009; NEF, 2009; Clark and Senik, 2011; Deci and Ryan, 2006). Eudaimonic well-being goes beyond a respondent's reflective evaluation and emotional states to focus on functioning and the realisation of the person's potential. In developing the questionnaire on psychological well-being for the *European Social Survey*, for example, Huppert et al. (2009) characterise the "functioning" element of well-being as comprising autonomy, competence, interest in learning, goal orientation, sense of purpose, resilience, social engagement, caring and altruism. Eudaimonic conceptions of SWB thus differ significantly from the evaluative and affective components in that they are concerned with capabilities as much as with final outcomes. Because the measurement of eudaimonia identifies a central role for people's "needs" or "goals", the approach represents a useful response to the criticism that the measurement of SWB is built purely on hedonistic philosophy, and also aligns itself with many people's perceptions of what is important to value in life.

While a consensus has emerged regarding the distinction between life evaluation and affect, the conceptual structure of eudaimonic well-being is less well fleshed out. It is not clear, for example, whether eudaimonic well-being describes a uni-dimensional concept in the sense of life evaluation, or a range of different concepts. It is clear, however, that eudaimonic measures capture important aspects of people's subjective perceptions about their own well-being that are not covered by either life evaluations or affect. For example, having children has a negligible (or even mildly negative) correlation with average levels of life evaluation (Dolan, Peasgood, and White, 2008), and child care (even of one's own children) is associated with relatively low levels of positive affect (Kahneman et al., 2004). This conflicts with the intuitive assumption that children, at least for those who choose to have them, contribute in some way to their parent's well-being. Indeed, people with children report higher average levels of meaning or purpose in their lives than other respondents (Thompson and Marks, 2008).

Concepts of "worthwhileness" or purpose appear crucial for understanding (and even predicting) behaviour, specifically why and when people engage in various activities or how they make decisions affecting their life course. White and Dolan (2009), for example, use day reconstructions to measure rewards associated with various daily activities. They find discrepancies between activities that people find pleasurable versus those found to be rewarding or meaningful. As noted above, activity based data indicate that time spent with children is relatively more rewarding than pleasurable, whereas time spent watching television is relatively more pleasurable than rewarding (NAS, 2013, p. 19). Similarly, people do many things that are pleasant even if they are not viewed as having much long-

term meaning or positive impact on future well-being. Either the pleasure or the purpose may be drivers of behaviour (Kahneman and Krueger, 2006).

While there is less agreement on the appropriate strategy to adopt when measuring eudaimonia than is the case for life evaluations or experienced well-being, three different measurement approaches have emerged: Economists have focused on meaning and purpose as the element of eudaimonia that most clearly captures additional information to other dimensions of SWB and which can be clearly distinguished from personality (Dolan, Layard, and Metcalfe, 2011). This has been reflected in the inclusion of a single question on meaning and purpose in the measures of SWB collected by the UK Office for National Statistics. An alternative strategy is that adopted in the well-being module of the European Social Survey (Huppert et al., 2009) where a battery of questions relating to different aspects of psychological well-being is collected, allowing for an analysis of the different concepts that are grouped together as eudaimonia. Finally, an extensive literature has emerged on a measurement of well-being grounded in mental health promotion. Mental well-being incorporates many of the elements of eudaimonia, but also combines these with measures of life evaluation and experienced well-being to provide a single index of overall psychological flourishing. A good example is the Warwick-Edinburgh Mental Well-being Scale (Tennant et al., 2007).

2.4. Relationships among the dimensions of SWB

While life evaluation, experienced, and eudaimonic well-being are all conceptually distinct, it is helpful to understand how they relate to one other. In the SWB continuum identified above, one end is demarcated by a point-in-time reference period and is purely hedonic ("How do you feel at this moment?") while the other involves evaluation of a comparatively very long reference period ("Taking all things together, how would you evaluate your life?"). Momentary assessments of affect represent the shortest framing period while global assessments of affect over the past day or even several weeks are at the longer end for experience measures. As the reference and recall periods lengthen, a measure is less dominated by actual experience and is more influenced by personality and/or cognitive reflection. Specification of the reference period has a determinative impact on the results of a survey and, indeed, on what nature of what is being measured (NAS, 2013, p. 29).

Figure 1 below provides a simple model of the different elements of a SWB measurement framework. The model emphasises three dimensions involved in the measurement of SWB. These are: (1) the measurement concept; (2) the sub-components of well-being; and (3) their determinants. The list of determinants and sub-components in the figure is illustrative rather than exhaustive − the model is intended to serve as an organising framework for thinking about the scope of SWB.

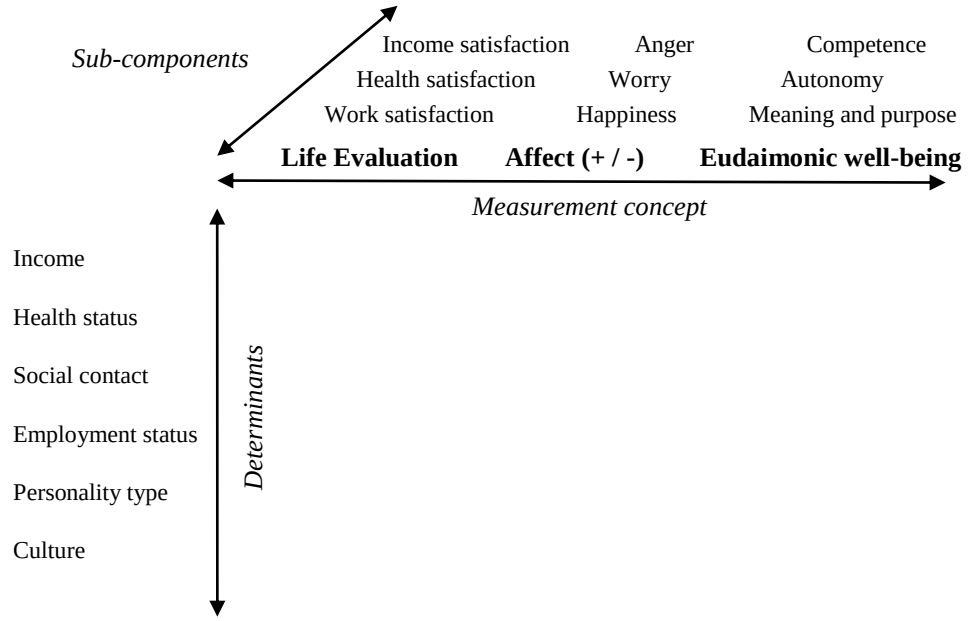

Figure 1. A simple model of subjective well-being
Source: OECD, 2013.

Empirically, there is extensive evidence about the relationship between measures of affect and overall measures of life evaluation. Diener, Kahneman, Tov, and Arora (in Diener, Helliwell, and Kahneman, 2010) reveal a high correlation (0.82) across countries between the most commonly used average measures of life evaluation, but a much lower correlation (0.55-0.62) between average affect balance and either of two life evaluation measures (life satisfaction and the Cantril Ladder). Similarly, at the individual level, Kahneman and Krueger (2006) report only a moderate correlation (0.38) between life satisfaction (an evaluative measure) and net affect.

Above, we have already hinted at how eudaimonia relates to experienced and evaluative well-being. For measurement, it may not make much difference whether sense of purpose contributes directly to positive or negative emotions or is positioned alongside but separate from them as a distinct sentiment. What matters is that the adjectives for purpose (e.g. fulfilment) are distinct from those used for pleasure (e.g. fun) and that a range of good feelings, emotions, or sentiments contributes to overall well-being.

A body of evidence exists on the empirical relationship between eudaimonic well-being and other dimensions of SWB suggesting that the correlation is smaller than is the case between affect and life evaluations. Clarke and Senik (2011) report a correlation between life satisfaction and four different aspects of eudaimonic well-being of between 0.25 and 0.29. Diener et al. (2009) report a correlation of 0.62 (N=563, p< .001) between their Psychological Well-Being

Scale and the evaluative Satisfaction with Life Scale, and correlations of 0.62 and 0.51 respectively between the Psychological Well-Being Scale and the positive and negative subscales of the Scale of Positive and Negative Experience (N=563, $p < .001$ in all cases). Huppert and So (2009) found a correlation of 0.32 between flourishing and life satisfaction in European Social Survey data. Among the European Social Survey sample overall, 12.2% met the criteria for flourishing, and 17.7% met the criteria for high life satisfaction, but the percentage for both flourishing and high life satisfaction was 7.2%.

Table 1 below gives the correlations between individual measures of life evaluation derived from the Gallup World Poll (life satisfaction), positive affect, negative affect and eudaimonic well-being (purpose) across 362 000 respondents in 34 OECD countries. The correlation is highest between the two measures of affect, at -0.3855, and lowest between purpose and negative affect, at -0.091. Life satisfaction has a correlation of about 0.23 with both measures of affect, and of 0.13 with purpose. While all the coefficients in Table 1 show the expected sign and all are significant at the 0.1% level, none of the measures have a correlation near 1, indicating that the different measures capture different underlying phenomena.

Table 1. Correlation coefficients among purpose, life satisfaction, positive affect, and negative affect at the individual level, 2006-2010

	Purpose	Life satisfaction	Positive affect	Negative affect
Purpose	1.000			
Life Satisfaction	0.134	1.000		
Positive Affect	0.142	0.229	1.000	
Negative Affect	-0.091	-0.231	-0.3855	1.000

Note: The precise measures used are the so-called "Cantril Ladder" for life satisfaction, an "important purpose" in life for purpose, and the sum of "yes" responses to smiled yesterday, experienced joy yesterday, and was well rested yesterday for positive affect and an equivalent index based on experience of sadness, worry and depression for negative affect.

Source: Gallup World Poll.

Because dimensions of SWB are distinct, and cover different reference periods, they can go in different directions. For example, studying hard for years to become a surgeon or working in devastated areas of the globe to alleviate poverty may not be immensely pleasurable but may ultimately yield high life satisfaction or reported sense of purpose. Individuals who have a longer-term focus and are more "achievement oriented," may at times sacrifice daily experience for longer term objectives and anticipated life satisfaction in the

future. The fact that people exhibit high and low discount rates means that they do not all have same focus. Individuals who focus primarily on daily experiences—due to low expectations, lack of agency, or imposed social norms—may have less incentive to invest in the future.

Relative to life satisfaction, experienced well-being is more directly related to the environment and context of people's lives. Using data from the Gallup World Poll, Deaton (2012) found, for example, that health state correlates more strongly with experienced well-being (though it is also important for evaluative well-being) as are marital status and social time (see also Boarini et. al., 2012). Other aspects of daily behaviour, such as the nature of a person's commute to work and the nature of a person's social networks, are reflected in positive and negative affective states (separable aspects of experienced well-being). The quality of people's daily experiences is also linked to health status and other outcomes via channels such as worry and stress on the one hand and pleasure and enjoyment on the other. Evaluative well-being, while also sometimes influenced by these factors, is more likely to reflect people's longer-term outlook about their lives as a whole. It may also be related to, and reflected in, longer-term behaviours such as investments in health and education. These distinctions make experience measures ideal for assessing emotions as they fluctuate from moment to moment and in response to day-to-day events and activities. In contrast, life satisfaction is more likely to reflect general, long-lasting factors such as unemployment, income, or a happy marriage, although it is easy to see how these circumstances could directly impact emotions on a day to day basis as well (NAS, 2013, p. 92).

These nuances and interactions led the NAS panel to conclude that: "To make well-informed policy decisions, data are needed on both [experienced well-being] and evaluative well-being. Considering only one or the other could lead to a distorted conception of the relationship between SWB and the issues it is capable of informing, a truncated basis for predicting peoples' behaviour and choices, and ultimately compromised policy prescriptions" (p. 93). A similar view is expressed in the *OECD Guidelines on Measuring Subjective Well-being*, which recommends that measures of affect and eudaimonia be collected alongside measures of life evaluation because they capture different aspects of SWB (with a different set of drivers) and because the different measures are affected in different ways by cultural and other sources of measurement error.

3. The role of national statistics in SWB measurement

3.1. Principles of official statistics

Official statistics are produced to meet the needs of policy-makers in planning and assessing the impact of policy decisions, and to inform the general public about the state of society. Academics and the media are also important users of official statistics, contributing to a better understanding of society and informing the public and decision-makers.

The principles of official statistics generally reflect the view that information is collected only when there is good reason and for a clear purpose. The OECD framework for data quality identifies relevance as the first of the seven key dimensions of quality. Relevance implies that the value of data "is characterised by the degree to which that data serves to address the purposes for which they are sought by users" (OECD, 2013). Similarly, the *United Nations Fundamental Principles of Official Statistics* asserts that the role of official statistical agencies is to compile and make available "official statistics that meet the test of practical utility… to honour citizens' entitlement to public information."

There are sound ethical and practical reasons why official statistical agencies insist on having a clear understanding of the uses of any proposed statistical measures. Many official statistical agencies have the power to compel responses from respondents. That is, respondents are legally required to provide information when approached by a national statistical agency. The corollary of such authority is the requirement for national statistical offices to use data responsibly. From an ethical standpoint, only information that is sufficiently important to justify the intrusion into respondents' lives should be collected. The International Statistical Institute's *Guidelines on Professional Ethics* notes that:

Statisticians should be aware of the intrusive potential of some of their work. They have no special entitlement to study all phenomena.

Over and above this ethical concern is a practical one. Even if compliance is legally mandated, the quality of resultant data depends heavily on preserving a good relationship between respondents and the official statistical agency. This is undermined if the statistical agency cannot articulate why the data being collected is important and how it will be used. Additionally, statistical agencies must be careful not to over-burden respondents and jeopardise the good will on which high-quality responses depend. Because of this, collecting measures of SWB will have an opportunity cost in terms of other data that will not be collected in order to produce such measures. If SWB measures are to be included in official statistics, therefore, it is essential to be clear about how they will be used.

3.2. Comparative advantages/disadvantages of government surveys

The fact that NSOs have historically led the way in the development of population surveys—both general and specialized (e.g. health interview surveys, time use, neighbourhood environment) − for research purposes puts them at a comparative advantage for collecting data on some dimensions of SWB, and perhaps at a comparative disadvantage for others. Traditional government surveys work especially well for large, cross-sectional formats. Life satisfaction or global yesterday questions of the type developed by the UK ONS are easily incorporated. Cross-sectional surveys are most often used to address group differences −for example, in the SWB context, are older people happier than younger people? Are females more stressed than males? Or, do people in high income countries report higher life satisfaction than those in low income

countries? A prime attraction of including SWB questions in large government surveys is their ability to accurately detect these group differences in a minimally burdensome way.

Among the key strengths of large scale surveys run by national statistical agencies are large sample sizes, high response rates, and ability to spread the enumeration out over a long period of time. These factors are important as they compensate for the weaknesses inherent in many research surveys, including those forming the basis of much of the SWB literature such as the *World Values Survey* and the *Gallup World Poll*. In both these cases, small national samples and short periods of enumeration make the error terms associated with SWB measures large, and raise the risk that transient events such as the weather, holidays, or news stories will impact on how people respond in unintended ways.[7] By way of contrast, a large sample size reduces the error term associated with SWB measures and a long (ideally annual) enumeration period will largely eliminate measurement bias due to one-off events.

More generally, surveys carried out by government statistical agencies generally collect higher quality information on potential covariates – such as income, labour force status, or education – than is possible in smaller unofficial surveys. Because the quality of SWB analysis depends not only on the quality of the SWB measure, but also on the quality of the other measures used in the analysis, surveys from national statistical agencies offer the opportunity for analysis not possible otherwise. For example, the lack of high quality income measures in surveys that include subjective well-being questions has been a factor limiting research in a number of areas. The relationship between income and subjective well-being has been a subject of interest since 1974 when Richard Easterlin identified the so-called "Easterlin paradox": that higher income is associated with higher happiness both between individuals and across countries, but there is no evidence that average happiness increases as average income increases over time (Easterlin, 1974). Understanding the causes of the Easterlin paradox is a high research priority because of the implications the paradox has for a range of policies. On a more technical level, one of the main policy uses for measures of subjective well-being is estimating the value of non-market outcomes. This involves obtaining precise measures of the impact of people's own income on their subjective well-being and comparing this to the impact of marginal change in the non-market outcome in question on subjective well-being.

For both better understanding the Easterlin paradox and estimating the value of non-market outcomes, the quality of income measures in surveys is at least as important as the quality of subjective well-being measures. While national statistical offices collect high quality information on household income, and are increasingly collecting measures of subjective well-being, there are currently few data sources that bring the two together. Those surveys – both official and non-

[7] For some measurement objectives, usually associated with experienced well-being, it is appropriate to factor in such influences. For life evaluation, it typically is not.

official – that include measures of subjective well-being typically collect income only in broad bands, and in the case of non-official surveys, often also have very high item-specific non-response rates for the income question. Filling this gap is a priority for the near future.

Another key strength of collecting data through large scale government surveys is the ability to conduct high-quality experimental trials to establish the impact of different methodological issues. For example, a split-sample randomized trial using experimental national data conducted by the UK Office for National Statistics (ONS) reported an effect of question order on multiple-item positive and negative emotion questions (Office for National Statistics, 2011). Asking negative emotion questions first produced lower scores on some positive emotion items for the adjectives "relaxed," "calm," "excited," and "energized." When positive emotion questions were asked first, the mean ratings for negative emotion questions were generally *higher*−except in the case of "pain"—and the increase was statistically significant for the adjectives "worried" and "bored" (OECD, 2013, p. 87). Similarly, when the order of positive and negative adjectives was varied, Krueger et al. (2009) observed higher ratings of positive emotions in a positive-to-negative order and lower ratings of negative emotions in a negative-to-positive order.

3.3. Measurement hurdles

Measuring SWB faces a range of methodological challenges. Almost all of these are shared with other survey measures, including those of notionally "objective" outcomes, but there is reason to believe that some of these issues may be more significant for SWB measures than for many other subjects[8]. Among these are context effects (such as the weather at the time of the interview, sports news on the day), framing effects (such as question order), mode effects (how the survey was carried out) and potential cultural bias[9]. These factors can affect answers to questions on life evaluation, affect, or eudaimonia.

Survey Mode is also a major methodological concern (see OECD, pp. 102-108). Dolan and Kavetsos (2012) investigated the differences between interviewer-administered and telephone-administered responses to the UK Annual Population Survey. The authors examined (a) the impact of survey mode on SWB reports and (b) the determinants of SWB by mode, using the April-September 2011 pre-release of the survey data. Their analysis found large differences by survey mode; in fact, mode effects in the data swamped all other effects.

Although the methodological challenges associated with collecting information on SWB are real, it is important not to overstate them. For some

[8] While the range of measurement issues – to do with survey context, question ordering, survey mode and many other factors – are briefly touched on in this paper, they are discussed in greater detail in papers by Lucy Tinkler and Paul Allin in this volume.

[9] For a comprehensive review, see Schwarz and Strack (1999).

questions, researchers may want to include the influence of context. For example, when looking at experienced well-being while using public parks, the context is not a contamination of the measure: it is the object of study. In some cases buffer and transition questions that precede and follow SWB question modules may help reduce or eliminate context or framing effects. For example, Deaton (2010) shows how including a buffer or transition question between political questions and life-evaluation questions largely eliminates a previously detected item-order effect. More generally, when the goal is to draw conclusions about a population, only influences that affect the sample but not the population as a whole, undermine the purpose of assessment. Provided that surveys are conducted in a consistent fashion (i.e., without changes to mode or question wording) and are enumerated over a long period of time, these issues are not generally significant.

Cultural bias is potentially a more difficult form of error to address. Taken in a general form, cultural bias can be thought of as differing response styles across different groups in the population of interest. In this case, even use of the same survey methodology at the same time will not eliminate sources of bias. A particular concern in this respect is the comparison of average levels of SWB between countries as there is prima facie evidence that response styles do vary between countries (OECD, 2013), and this will have an important impact on the inferences drawn from the data.

Several special challenges arise when measuring experienced well-being. A number of national and international surveys have used single-day assessments to measure experienced well-being−that is, assessments that target affect or broader experience for a single day. In the US, for example, the Health and Retirement Study, the Disability and Use of Time supplement to the Panel Study of Income Dynamics, and the Gallup-Healthways survey employ single-day hedonic assessments; as do the English Longitudinal Survey of Ageing and the surveys on well-being of the UK Office for National Statistics (ONS) (NAS, 2013, p. 52). Typically, these surveys ask respondents about their experiences from the previous day. The NAS (p. 55) report concluded:

> Global-yesterday measures represent a practical methodology for use in large population surveys. Data from such surveys have yielded important insights−for example, about the relationships between experienced well-being and income, age, health status, employment status, and other social and demographic characteristics. Research using these data has also revealed how these relationships differ from those associated with measures of evaluative well-being. Even so, there is much still to be learned about single-day measures.

One practical limitation of *end-of-day* − as opposed to global yesterday, which are often the default for large surveys) measures, and a reason that they have not been used more by statistical agencies − is that large population surveys often depend on telephone interviews conducted throughout the day, not just at the end

of the day. Because of the survey timing requirement, end-of-day instruments have typically been less practical for use in general surveys. However, newer technologies, such as use of interactive mobile phone assessments, may offer solutions to some of the data collection constraints associated with end-of-day methods (NAS, 2013, p. 53). There has been little systematic research into how the recall and contextual influences act differentially between end-of-day and global-yesterday measures, and how well either correlates with averages from momentary readings[10].

Additionally, global-yesterday measures are limited in terms of creating a more detailed understanding of the drivers of experienced well-being over the course of the day (e.g. variation at the individual level). For this level of analysis, momentary assessments or, at the least, time-use or activities-based data − for example, data generated by day reconstruction methods (DRM) − are needed (NAS, 2013, p. 55). For some research and policy questions, contextual information about activities engaged in specific behaviours and proximate determinants is essential. For example, to investigate how people feel during job search activities, while undergoing medical procedures, or when engaged in child care, something more detailed than a global daily assessment is needed. Activity based measures attempt to fill this measurement need (NAS, 2013, p. 59).

An attractive feature of DRM is its capacity to combine time-use information with the measurement of affective experiences. Capturing the time-use and activity details of survey respondents enhances the policy relevance of experienced well-being measures by embedding information about relationships between emotional states and specific activities of daily life (NAS, 2013, p. 66). Additionally, for large surveys, DRM can be administered with less intrusion and lower burden than momentary assessment tools while still gathering fairly rich and detailed information. By asking participants to first recall the events of their day and then provide ratings associated with them, DRM exploits the fact that, while memories of ongoing experiences such as pain and mood are flawed, Memory for discrete events is more accurate (Robinson and Clore, 2002) (NAS, 2013, p. 60). For some questions (e.g. predicting consumer behaviour or whether or not a person is likely to repeat a medical procedure), a reconstructed assessment of experienced well-being may be more relevant than EMA; it may also be better at predicting a policy's impact on people's choices, but worse at assessing a policy's impact on experience.

This kind of data collection has already been successfully developed by statistical agencies. In the United States, ATUS has, since 2010, included a module asking respondents about feelings (pain, happiness, stress, sadness, tiredness) during specific episodes of the day. The ATUS SWB module is an

[10] Though Christodoulou et al. (2013) compared to DRM – see Dylan M. Smith in this volume. The validity of different measures addressed in a paper by Paul Allin in this volume.

abbreviated version of a DRM approach (NAS, 2013, p. 23). Regarding the DRM, the NAS report concluded:

> Preliminary assessment of DRM measures of mood and physical symptoms suggests that they reasonably approximate summary measures created from EMA protocols. An attractive feature for survey objectives is that the DRM approach goes beyond simply addressing who in the surveyed population is happy to identifying when they are happy. Additionally, it appears that the DRM is less burdensome on respondents than experience sampling, and it might reduce memory biases that are inherent in global recall of feelings. The DRM is thus a promising method for assessing feelings, mood, and physical symptoms that accompany situations and activities more efficiently than with EMA methods and with greater specificity and accuracy than traditional recall-based methods (NRC, 2013, p. 63) [11].

Similarly, INSEE (the French national statistical agency) has collected data on experienced well-being through the French time use survey – the *Enquete Emploi du Temps 2010*. This survey used a different approach to the DRM strategy adopted by the ATUS SWB module. Rather than collecting detailed information on multiple different affective states for just three episodes in each diary day, the *Enquete Emploi du Temps* requires respondents to rate each activity in the time use diary on a 7 point scale from very unpleasant to very pleasant. This collects a far more comprehensive picture of the activities sampled at the price of less detail on each activity. The relative strengths and weaknesses of the French and American approaches is an area for further research.

For obvious reasons, surveys do not work easily for momentary assessment. The required instruments are difficult to scale up to nationally representative surveys and impose a high respondent burden. This said, monitoring and survey technologies are changing rapidly and the ways in which government agencies administer surveys will surely evolve alongside and new measurement opportunities will come on line. Considered in terms of comparative respondent burden, it may become less intrusive to respond to a smartphone beep than to fill out a long-form survey. So, while EMA may not now be practical for flagship population surveys, real-time analyses may become so. As technology advances, such modes could become feasible, even for large-scale surveys at reasonable cost. Large-scale (more general) surveys could build in the possibility of mapping the data from single-day measures with the data from more detailed studies for a subset of the sample. Experiences in real time, because they are especially relevant to health, have been incorporated into health examination surveys, so there is precedent. It is also possible to monitor blood pressure and other physical signals related to affect in real time (NAS, 2013, p. 51).

[11] Smith, et al. (2012). A Test Comparison of EMA and DRM estimates supports above conclusion.

3.4. Current state of play

When the report of the Commission on the Measurement of Economic Performance and Social Progress (Stiglitz et al.) was released in 2009, the national statistical offices of only three OECD countries collected regular measures of life evaluation (Canada, Israel, New Zealand) and, of these, only Statistics Canada conformed to current best practice. No OECD country was regularly collecting measures of affect or eudaimonia. In the six years since then, this situation has transformed. Among the 34 OECD countries, 32 now collect measures of SWB – mostly life evaluation – through their national statistical agencies in a way that is broadly comparable.

Table 2. Subjective well-being measures in official statistics

EU-SILC coverage (2013 ad-hoc module; life sat, affect and eudaimonia – freq. tbc.)		**European countries with additional collections**	**Other OECD countries with SWB measurement initiatives**	**OECD countries with no current NSO data collection**
OECD countries (25): Austria Belgium Czech Republic Denmark Germany Estonia Finland Greece Ireland Spain France Iceland Italy Luxembourg Hungary Netherlands Norway Poland	Portugal Slovenia Slovakia Sweden Switzerland Turkey UK **Non-OECD countries (7):** Bulgaria Croatia Cyprus FYROM Latvia Lithuania Malta	• Austria (life sat, 2004-2012) • France (life sat*, in 2011; affect in 2010 time-use survey; freq. tbc.) • Italy (life sat*, from 2012, yearly) • Netherlands (life sat and happiness, from 1974) • Poland (life sat, in 2011; freq. tbc.) • UK (life sat*, affect* and eudaimonia*, from 2011, yearly)	• Australia (life sat**, from 2016; every 4 years) • Canada (life sat*, from 1985; yearly) • Israel (life sat, from 2006; life sat*, affect* and eudaimonia* from 2013, yearly) • Korea (life sat*, affect* and eudaimonia*, from 2013, yearly) • Mexico (life sat* and affect, in 2012, freq, tbc.) • New Zealand (life sat** and eudaimonia**, from 2014, every 2 years) • United States (affect and experienced eudaimonia in 2011 time-use survey; freq. tbc.)	• Chile • Japan

**Questions broadly in line with OECD Guidelines; **Questions planned to be in line with OECD Guidelines*

The largest data collection exercise is that of the UK ONS which, since April 2011, has included a set of four questions on the core of its Annual Population Survey (n=160,000) covering life evaluation, momentary emotional state, and worthwhileness:

- Overall, how satisfied are you with your life nowadays? [evaluative well-being]
- Overall, to what extent do you feel the things you do in your life are worthwhile? [eudaimonic well-being]
- Overall, how happy did you feel yesterday? [experienced well-being]
- Overall, how anxious did you feel yesterday? [experienced well-being]

One of the most important developments identified in Table 2 is the European Union inclusion of a well-being module as an add-on to the main EU survey of

living conditions (EU-SILC). This module includes a question on life evaluation directly comparable to the OECD primary measure and a eudaimonic question that is very close to the one in the OECD core measures. As EU-SILC covers 27 EU countries as well as Croatia, Iceland, Norway, Switzerland, and Turkey, this extends the available data to the majority of the OECD, albeit with data updated only when the well-being module is run every six years. More importantly, although the decision has not been finalised, Eurostat (the EU Statistical Agency) has indicated that it is also considering including the primary life evaluation measure in the core of EU-SILC from 2020. This will make high quality annual data on life satisfaction available for the majority of the OECD on an ongoing basis.

3.5. Strategies and priorities

The nature of the policy or research question being asked dictates the appropriate SWB construct to measure and may suggest an approach to data collection. For example, if the dimension of interest is known to be sensitive on a very short time frame and responds to daily activities and events but is somewhat stable over long periods, a cross-sectional data collection conducted every 2 years may not be useful. In such cases, a high-frequency approach (even if it involves a much smaller sample) might be most informative. Similarly, if a measure varies a great deal from individual to individual on a given day but does not react very much to exogenous events (financial shocks, changes in employment rates, etc.) and tends to wash out at high aggregate levels, it may not be a particularly insightful construct to track at national levels over time (NAS, 2013, p. 16).

The Commission on the Measurement of Economic Performance and Social Progress (Stiglitz et al., 2009) concluded that, where feasible, inclusion of SWB questions on the largest population surveys will produce useful information. However, because it would be beneficial to have information about different sets of covariates for different applications, it is unlikely that an identical module could be simply plugged into different surveys to suit the many envisioned purposes for SWB data. If harmonized modules were developed that were short enough, they could in principle be included in a range of surveys. However, for surveys with a specific orientation (e.g. understanding the conditions of retirees or the time use of individuals) it would typically be preferable to tailor questions to research objectives. For example, the CPS (in which the American Time Use Survey module resides) is designed to optimize employment measures at specific levels of geographical specificity.

This diversity in the research landscape in which SWB is relevant suggests a multidimensional approach to data collection. Large-scale population surveys−such as the four-question module in the UK Integrated Household Survey or the Gallup World Poll−make up one component of a comprehensive measurement program. Data from these surveys, typically drawn from global-yesterday measures of experienced well-being and from life-evaluation questions,

provide the large sample sizes essential for repeated cross-sectional analyses capable of identifying and tracking suffering or thriving subgroups and for research on special populations such as the unemployed for whom life expectancy is falling.

The second prong of a comprehensive measurement program is smaller or more specialized data collections. One option is to construct experiments or pilots within existing large survey programs. The advantage of targeted studies and experimental modules is that they can be tailored to address specific questions of interest to researchers and policy makers−whether about health care, social connectedness of the elderly, city planning, airport noise management, or environmental monitoring.

The third prong of an ideal data infrastructure would consist of panel studies designed to document changes in SWB over time. The inclusion of SWB measures in *Understanding Society* (formerly the *British Household Panel Survey*) and the *German Socio-economic Panel* has already contributed greatly to understanding issues relating to both causality and to adaptation over time (e.g. Lucas, 2007; Lucas et al., 2004). How individuals' experienced well-being and life satisfaction change over time and in reaction to events and life circumstances cannot be fully understood without longitudinal information, which may also help to make progress on causality questions (e.g. does getting married make people happier, or are happier people more likely to get married?). The policy relevance of monitoring SWB changes over time is clear where, for example, it is important to know the full impact on people of new legislation or on outcomes of experiments such as the Oregon Health Care Study (NAS, 2013, p. 107). Schuller et al. (2012) reviews the contribution of longitudinal data in analyzing SWB responses for a range of key well-being domains, such as relationships, health, and personal finance.

A final prong of an ideal data collection is information on experienced well-being. As described above, momentary sampling methods have been central to SWB research but largely out of practical reach for adoption by national statistical offices. However, rapid changes in technology and in the way the public exchanges information have brought the world to a point where momentary assessment techniques may now be on the horizon for national statistics. Regardless of developments in EMA, collecting experienced well-being data through a DRM approach in nationally representative time use surveys has been demonstrated to be feasible both through the *ATUS* and the *Enquete Emploi du Temps*. The UNECE Guidelines for Harmonising Time Use Surveys (UNECE, 2014) and the *OECD Guidelines on Measuring Subjective Well-being* both recommend that national statistical offices should move to collect experienced well-being data in time use surveys. Recently Statistics Canada has become the first national statistical office to move in this direction following the examples of the Bureau of Labor Statistics in the USA and INSEE in France. Precisely knowing how people are doing emotionally and what they are doing in the

moment can shed light on the effects of commuting, air pollution, child care, and a long list of areas with clear ties (NAS, 2013, p. 108).

The Mappiness project (mappiness.org.uk), designed to investigate well-being effects to the public associated with open green space in the London area, allows monitors to look at individual-level variation for people located in different outdoor environments. This project provides a clear example of the emerging methods to capture SWB in the context of EMA measures and the role of portable recording−in this case the use of cellphones and global positioning system (GPS) tracking. The British Millenium Cohort Study is considering use of geospatial cellphone responses as a post-survey supplement. There are still major unresolved data quality and representativeness issues in this world of new data and big data. For instance, the sampling properties are largely unknown for data generated by social media, phone records, Internet usage, and the like. Much more will need to be learned about distributional characteristics of various underlying subpopulations.

Social media data and other kinds of unstructured data (those, such as administrative records or company-maintained information, produced initially as a by-product of non-statistical purposes) may become increasingly useful for shedding light on trends in people's emotional states. Word mining exercises have been used to show patterns in emotional states—for example, a Facebook happiness index showed the standard weekend and holiday effects and expected changes associated with major events, such as disasters. The words people use on social media such as Twitter, Facebook, and Google search queries are a rich, if imperfect, source of information about their personality and psychological state. Additionally, analyses of data generated by social media and other Internet activities will produce insights relevant to public policy (see the discussion below of relevance to understanding social or political movements such as the Arab Spring).

4. How do different dimensions of SWB link with policies?

Informing policy−or at least the potential to do so−is a critical criterion for deciding whether it is worth the time and cost of measuring SWB in national flagship population surveys or in more focused domain-specific surveys. It is clear that different kinds of SWB measures inform different kinds of policies. For example, optimizing end-of-life care decisions may give greater weight to short-term concerns−minimizing day to day suffering−and therefore suggest a need for experience based measures. Education and employment policies may focus more on life satisfaction or even eudaimonic concerns, for which evaluative measures are highly relevant. In either case, assessment is needed about the extent to which SWB adds analytic content beyond the existing "objective statistics" such as those we have come to rely on in such research and policy areas as poverty (e.g. income data) and health (e.g. vital statistics).

The unique policy value of SWB measures may lie not in assessing how income or other variables relate to an aggregate-level tracking of emotional states or life satisfaction, but in discovering actionable relationships that might otherwise escape attention in order to better understand the full of impact of commuting patterns, accessibility of child care, exercise programs, interaction and connectedness with neighbours and friends, the presence of neighbourhood amenities and other city planning issues, divorce and child custody laws, and the like (NAS, 2013, pp. 88-89).

The intended use for measures of SWB also affects judgements about the validity of such measures. In the remainder of this section, we outline the major uses of SWB measures: (1) complementing objective measures of the economy, health, and society; (2) to better understand the drivers of well-being at the level of the individual; (3) for policy evaluation and cost benefit analyses; and (4) for identifying potential policy issues.

4.1. The role of SWB as a complement to objective economic, health, and social measures

SWB measures offer significant potential for complementing conventional economic, social, and health metrics by providing an alternative yardstick of progress that is grounded in people's experiences or evaluations. Traditional market-based measures alone cannot provide an adequate portrayal of quality of life, which suggests a need to shift some portion of the measurement focus from economic production toward people's well-being. The underlying argument is that national policies should better balance growth in market production with nonmarket dimensions of well-being that cannot be captured well by conventional measures. In particular, being grounded in peoples' experiences and judgements on multiple aspects of their life, SWB measures provide information about the net impact of changes in social and economic conditions on the perceived well-being of respondents, reflecting differences in tastes and preferences among individuals. An example of how these measures can change perceptions about progress is provided by Box 4.1, in respect of the "Arab Spring."

In addition to information on aggregate trends, SWB measures can also provide a picture of which groups in society are most (dis)satisfied or experience the best or worst lives that reflect, among other things, the impact of tastes, aspirations, and life circumstances. Migrants, for example, may be more motivated than the rest of the population by income relative to other factors (Bartram, 2010), as this is a primary motive for their decision to move abroad. This heterogeneity makes assessing overall migrant well-being compared to the rest of the population challenging. However, because SWB measures incorporate the impact of different weights that people attach to aspects of their quality of life, they have the potential to add an important dimension to analyses in situations involving comparisons between population groups.

Box 4.1. Subjective well-being, GDP growth and the "Arab Spring"

For policy-makers, measures of SWB are valuable as indicators of progress when they can alert them to issues that other social and economic indicators might fail to identify. One recent example where measures of SWB demonstrated their ability to capture important elements of well-being not captured by more traditional measures was the decline in country-average measures of SWB that occurred in Egypt and Tunisia in the years leading up to 2011, a decline that contrasts with the much more favourable evolution of GDP data. For example, Tunisian real GDP per capita increased from USD 8,891 in 2008 to USD 9,489 in 2010, a real gain of around 7%. However, the proportion of the population indicating a high level of satisfaction with their life as a whole fell from 24% to 14% over the same period (Gallup, 2011). Egypt showed a similar pattern from 2005 to 2010, with a real gain in GDP per capita of around 34% and a decline in the share of respondents classified as "thriving" by almost half. This illustrates how subjective perceptions can provide information on very significant outcomes in societies that other conventional indicators such as GDP growth do not provide.

An additional use of SWB measures is for monitoring progress in aggregate cross-country comparisons, such as those included in *How's Life?* (OECD, 2011). Because controlled experiments are typically impossible, cross-country comparisons of SWB outcomes are one way to learn about the strengths and weaknesses of different policies. When SWB measures are sensitive to a different range of drivers than are other social and economic indicators, they provide additional information about the consequences of a particular policy. A crucial issue in using SWB in this way, however, is the degree to which cross-cultural comparisons of such measures are valid.

Interest by the general public and the media in using measures of SWB as complements of measures of progress represents another valid rationale for public data collection. Of particular interest to these users is the question of whether things are getting better or worse overall, and for whom. As in the policy realm, SWB measures used for general public information purposes should be viewed as one set in the much broader array of indicators through which populations are monitored and insights about societal progress or deterioration are drawn.

4.2. The role of SWB in better understanding the drivers of people's well-being

A second major use of SWB measures is to contribute to a better understanding of the drivers of well-being at an individual level. If it can be established that SWB measures accurately capture the concepts that they claim to – an overall evaluation of life or the experienced moods and emotions of an individual over a period of time – they can be used to provide information about the relative contribution of different factors and circumstances to a person's well-being. The quality of the information will be tempered by measurement error and by the fact that a person's subjective perception of their well-being is not necessarily quite the same thing as their overall well-being (see Dolan, Peasgood,

and White, 2008; Helliwell and Wang, 2011; Boarini, Comola, Smith, Manchin, and De Keulenaer, 2012).

Subjective measures can be used to test specific hypotheses about what aspects of policy are most important to people. Halpern (2010), for example, refers to an instance in which the Merseyside police, in the United Kingdom, used data on how satisfied members of the public were with the service provided by the local force, alongside more traditional performance measures on crimes committed and offence resolutions. In contrast to the expected hypothesis – which was that minimising the response time from the police was of crucial importance for public satisfaction – the evidence showed that it was much more important that police arrived when they said they would. For minor issues not involving safety, what mattered was the punctuality rather than the speed of the response.

Going beyond simply identifying what matters to people, SWB measures can provide the basis for developing a better understanding of trade-offs when policy options involve comparisons of fundamentally different types of outcome (see box 4.2 below). Dolan and White (2007) note that this issue characterises many attempts to encourage "joined-up government," where costs and benefits of a particular intervention must be considered not just based on the outcome of concern to one agency, but also in terms of how choices affect the outcomes of other agencies.

Measures of SWB can potentially capture the combined effect on an individual's perception of their well-being of a range of different changes in life circumstances. For example, Ferrer-i-Carbonell and Frijters (2004) compare the magnitude of the impact of health satisfaction versus housing satisfaction on overall life satisfaction.[12] Similarly, Di Tella, MacCulloch, and Oswald (2003) investigate inflation, unemployment trade-off in terms of the effect on life satisfaction. While the so-called "misery index" weights the unemployment rate and inflation rate equally as indicators of the negative impact of macro-economic outcomes, Oswald and Maculloch's analysis suggests that the impact of unemployment on SWB is significantly greater than that of inflation.

Box 4.2. Using measures of subjective well-being to value life events

Measures of SWB provide a relatively straight-forward way of comparing the relative impact of fundamentally different life events in a quantitative way and, based on this, assigning such events a monetary value. Clark and Oswald (2002) present a method for valuing life events and, although the literature on using measures of SWB to value life events has expanded significantly since 2002, the basic methodology remains largely unchanged. Consider the results below from Boarini et al. (2012). The coefficients for the (base two) logarithm of household income, being married, and being unemployed are

[12] Consideration of initial sample variance in each measure is important here: if the sample has uniformly high levels of health satisfaction, but variable levels of housing satisfaction, housing satisfaction may look more important in a regression analysis, simply because it has more variation to associate with variation in the outcome measure.

shown, and express the change in life satisfaction (on a scale of 0 to 10) associated with a doubling of income, being married, or being unemployed, respectively, holding all else constant.

Event	Coefficient
Log Household Income	0.1482
Married	0.2584
Unemployed	-0.4643

Using these coefficients, it is possible to calculate the relative impact of being married compared to being unemployed on life satisfaction as 0.2584 / 0.4643 = 0.5565. Or, put more simply, being unemployed has almost twice the impact on life satisfaction as does being married.

Going beyond this, the monetary value of being married or being unemployed can be calculated by comparing the relevant coefficients with that associated with the coefficient for household income. Using the values presented above, the coefficient on being married is 0.2584 / 0.1482 = 1.7435 times larger than the impact of a doubling of household income. For a person with a household income equal to the average OECD per capita household disposable income ($17 286 at PPP, 2008), this is equivalent to 1.7435 x $17 286 = $30 138. For unemployment the comparable value is 2.930 x $17 286 = $50 647.

These values are intended to illustrate the techniques involved, and need to be treated with caution. In particular, it would be preferable to use panel data which might better capture a causal relationship (as do Clark and Oswald) rather than just correlation; potential biases in the data as well as appropriate model specification also must be evaluated (Fujiwara and Campbell, 2011).

4.3. The role of SWB in policy evaluation and cost benefit analyses

A third use of SWB measures is to assist in the evaluation of policies. This includes both the direct use of measures of SWB in formal policy evaluations as well as the more indirect – but possibly more important – role that they can play in cost-benefit analysis. For some initiatives – where the impact on subjective experiences of the population is the main object of the program – measures of SWB may even be suitable as the primary metric for assessing its success.

Many policy evaluations already include subjective measures of client satisfaction that gauge respondents' perceptions about what elements of a program are most valuable. More general measures of overall SWB, however, have some significant advantages over and above these more focused measures. Most importantly, measures of SWB provide information about the impact of an initiative on the respondent's SWB, rather than the impact that the respondent consciously identifies. These values can differ because people's judgements about the impact of a program may be influenced by their participation (i.e., they might be more prone to assign the cause of any recent changes in their well-being to the program rather than to other factors, knowing that this is what he/she is being asked about). Also, people may not be aware of all of the various feedback loops via which a policy programme affects them. For example, in evaluating an active

employment program, respondents might consider the direct effect on their well-being of both having a job and gaining additional income, but not the flow on well-being that would stem from changes in their time-use due to longer commuting. Because measures of SWB can capture the overall impact of a change on life circumstances, without requiring a cognitive judgement by the respondent on which causal pathways are being asked about, such measures provide useful additional information on the overall impact of a programme.

In some cases, measures of SWB can be better than conventional cost-benefit analysis at treating non-monetary outcomes. Examining the relative costs and benefits of a proposal is relatively straight-forward when the proposal is aimed at strictly economic outcomes, and the costs and benefits of the proposal can be obtained from the relevant market prices. However, where the aim of a proposal is to achieve outcomes that do not have an obvious market price, it is much more challenging to obtain meaningful values for analysing the relevant costs and benefits. Because much government policy is concerned with market failures, many government policies are correspondingly focused on achieving non-market outcomes.

The traditional economic approaches to cost-benefit analysis for non-market outcomes depend on either revealed preference or contingent valuation techniques to estimate "prices" for such outcomes. A revealed preference approach involves calculating values based on the shadow prices implied by observed behaviour, while contingent valuation techniques calculate values based on the "willingness to pay" for the outcome in question, as expressed by respondents to a hypothetical question in a survey. Clark and Oswald (2002) note that measures of SWB can provide the framework for such valuations by comparing the impact of a particular outcome on SWB with the impact of a change in income on SWB. By making such a comparison, it is possible to calculate the amount of money required to achieve the same increase or decrease in well-being as that caused by the outcome under assessment.

There is good reason to believe that, in some circumstances, measures of SWB have advantages over both revealed preference and contingent valuation for the purposes of cost-benefit analysis (see box 4.3 below). An obvious advantage is that many measures of SWB – such as overall life satisfaction – are relatively easy and cheap to collect. However, there are also more substantive methodological advantages that may be associated with using measures of SWB in this way. Revealed preference relies on strong assumptions about people's ability to know how an outcome will affect them in the future, and on the assumptions that markets are in equilibrium. Diener, Lucas, Schimmack, and Helliwell (2009) note that for market prices for houses to reflect the disutility of airport noise accurately would require that house purchasers are able to forecast how much the noise will impact them before buying the house. Similarly, in this example, it is difficult to disentangle the differences in house prices due to noise from differences in other aspects of house quality.

Box 4.3. *The Green Book* and life satisfaction

The Green Book is the formal guidance from the Treasury of the United Kingdom to other UK government agencies on how to appraise and evaluate policy proposals. The current edition of *The Green Book* dates to 2003, and provides advice on how officials should provide justification for a proposed government intervention, set objectives for the proposal, appraise the various options, and evaluate the effectiveness of the final action that results. In July 2011, *The Green Book* was updated to reflect the results of a review of valuation techniques for social cost-benefit analysis jointly commissioned by the Treasury and the Department for Work and Pensions (Fujiwara and Campbell, 2011). The review specifically focuses on the contribution that can be played by measures of SWB – particularly life satisfaction – alongside more traditional approaches to cost-benefit analysis. In summarising the conclusions of the review, *The Green Book* states (p. 58):

> *A newer, "subjective well-being approach" has been gaining currency in recent years. The "life satisfaction approach" looks at people's reported life satisfaction in surveys such as the ONS's Integrated Household Survey, which began including questions on respondents' subjective well-being in April 2011. The life satisfaction approach uses econometrics to estimate the life satisfaction provided by certain non-market goods, and coverts this into a monetary figure by combining it with an estimate of the effect of income on life satisfaction.*
>
> *At the moment, subjective well-being measurement remains an evolving methodology and existing valuations are not sufficiently accepted as robust enough for direct use in Social Cost-benefit Analysis. The technique is under development, however, and may soon be developed to the point where it can provide a reliable and accepted complement to the market based approaches outlined above. In the meantime, the technique will be important in ensuring that the full range of impacts of proposed policies are considered, and may provide added information about the relative value of non-market goods compared with each other, if not yet with market goods.*

While the amendment to *The Green Book* stops short of fully endorsing the use of life satisfaction measures for use in formally evaluating government programmes, the decision to make an interim amendment in itself signals strongly the importance that UK central agencies attach to obtaining improved measures of the value of non-market outcomes.

Contingent valuation also relies strongly on people's ability to make accurate judgements about how something will make them feel in the future. Dolan and Peasgood (2006) observe that people have difficulty imagining how good or bad different circumstances are actually going to be. Indeed, the "willingness to pay" surveys commonly used for contingent valuation are, to a large degree, measures of the SWB associated with a hypothetical scenario. Using measures of SWB to calculate the costs based on the actual impact of different life circumstances on SWB removes the hypothetical element from the equation. In addition, contingent valuation surveys tend to produce very different estimates of the value of outcomes for people at different points on the income distribution. This tends to result in either weighing the desires of the rich more heavily than the poor when assessing the costs and benefits associated with the proposal under consideration

or taking account of the marginal utility of income in calculating the final cost. The latter approach is difficult in the absence of robust estimates of the marginal utility of income (Dolan and White, 2007).

4.4. The role of SWB in identifying potential policy issues

An important feature of SWB measures is their ability to provide insights into human behaviour and decision-making. In particular, measures of SWB can help researchers better understand the difference between the *ex ante* beliefs that people hold about their future well-being (which form the basis for decisions) and the *ex post* outcomes that people achieve in terms of their SWB. A better understanding of these issues is important both for policy-makers and for the broader public. Policy-makers have an interest in understanding why people make the decisions that they do, because much public policy involves dealing with the consequences of systematic poor decision-making by individuals. Similarly, businesses and the general public have an interest in understanding how people's SWB shapes their behaviour.

One example of how subjective measures are useful to businesses and the broader public is the information they provide about the characteristics of good places to live and work and in turn how that predicts future behaviour. Clark (2001) has shown that measured job satisfaction predicts the probability of an employee leaving their job. Thus businesses might well have an interest in the measured job satisfaction of their employees and in understanding the determinants of job satisfaction.

Measures of SWB can also help shed light on various biases in the way people make decisions. Although people are generally able to predict whether events are likely to be pleasant or unpleasant, Wilson, Gilbert, and colleagues have described ways in which affective forecasting can be biased or faulty, particularly with regard to the intensity and duration of emotional reactions to future events (e.g. Wilson, Wheatley, Meyers, Gilbert, and Axsom, 2000; Wilson and Gilbert, 2006). Kahneman et al. (2006) show that people are prone to over-estimate the impact of income gains on their life satisfaction relative to other factors. Commuting, for example, has been found to have a strong negative impact on both measures of affect (Kahneman et al., 2006) and life evaluations (Frey and Stutzer, 2008). This suggests that people may be prone to over-estimating the positive impact of, for example, a new job with a higher salary but a longer commute.

There are also direct policy applications for better understanding the human decision-making process and the various biases and heuristics involved in it. Consider the case of policy options that incorporate a "default" option – for example, workplace retirement schemes that are set up on a basis of either "opt in" clauses, where a new employee does not join the scheme unless he/she ticks a box to join, or "opt out" clauses, where the reverse is the case. The fact that people respond differently depending on which default is selected – despite the fact that in neither case is there any compulsion – has raised policy interest in the

idea of "libertarian paternalism", which focuses on achieving better outcomes be setting policy defaults to influence people's behaviour in positive directions. Dolan and White (2007) note that information on SWB can be used to help set policy default options more optimally, by indicating which default options contribute most to SWB.

While a full accounting of SWB applications to research and policy is beyond scope here, the following examples hint at their diversity and potential (NAS, 2013, p. 89):

- Kahneman and Deaton (2010) and Stevenson and Wolfers (2013) used data collected in the Gallup-Healthways Well-Being Index to estimate the impact of income and income-normalized effects on life satisfaction and experienced well-being. Understanding the relationship could prove useful for informing tax and social program policies.
- Oswald and Wu (2009) used data from the Behavioral Risk Factor Surveillance System to rank the US states based on hedonic analyses of regional variation in such factors as precipitation, temperature, sunshine, environmental greenness, commuting time, air quality, and local taxes; all suggesting a role fir SWB data in assessing regional and city policies.
- Diener and Chan (2010) argue that people's emotional states causally affect their health and longevity, concluding that the data are compelling, though "not beyond a reasonable doubt" (NAS, 2013, pp. 87-88).
- Robert Sampson's Chicago neighborhoods study (Sampson and Graif, 2009) reveals the importance of connectedness to the well-being of neighbourhoods. One of many examples is the variation, even among relatively poor areas, in the resilience of different neighbourhoods to the 1994 heat wave in the city. Sampson's findings suggest the value of data on people's trust in neighbours, interactions, connectedness, as well as mechanisms whereby the built environment can promote SWB (though there is the alternative hypothesis being that happier people tend to have more autonomy over where they choose to live).
- Krueger and Mueller (2011) found that the SWB of the unemployed declines with the duration of unemployment spells; they also found that the time spent involved in job search is particularly unhappy and the unhappiness increases with the time spent in job search (measured both with life-satisfaction and sadness variables). These effects on the unemployed provide an example of how low experienced well-being related to the process could in the end undermine individuals' incentives to persist, ultimately reducing their capacity to achieve higher levels of evaluative well-being in the future.

From company policies that improve well-being − and possibly, in turn, improve productivity and lower absenteeism − to community or regional planning policies, SWB measures would appear most valuable when costs and benefits must be weighed in the absence of market or easily quantifiable elements.

Government consideration of spending to redirect an airport flight path to reduce noise pollution, funding alternative medical care treatments when more is at stake than maximizing life expectancy, or selecting between alternative recreational and other uses of environmental resources are possible examples.

Across these policy applications, experienced and evaluative dimensions of well-being may have very different implications (Diener, 2011; Graham, 2011; Kahneman et al., 2006). For example, actions aimed at enhancing longer-term opportunities may actually impart negative short term effects on daily experience. A policy designed to enhance living quality at the end of life, for example, focuses on the hedonic dimension (which is at least one of the objectives of palliative care, that is, relieving suffering), while a policy aimed at increasing educational opportunities of youth focuses on life evaluation (NAS, 2013, p. 91). Optimization of short-term versus long-term well-being (both at individual and aggregated levels) may imply different policy actions. A program to reduce fat intake or smoking may reduce experienced well-being in the short run but increase it (via the health covariate) over the long run.

The distinction between positive and negative affect and between suffering and happiness are also important with, arguably, minimization of the negative being more relevant to public policy.[13] The U.S. General Social Survey (GSS) registers exposure to negative circumstances and events experienced by people (e.g. hospitalization, death of a family member, eviction, crime victimization), and was designed to report "objective experiences that disrupt or threaten to disrupt an individual's usual activities, causing a substantial readjustment in that person's behavior" (Thoits, 1983). As described by Smith (2005), this approach has been used extensively not only to account for differing levels of reported well-being among individuals or groups but also for understanding and predicting individual illness (both psychological and physiological); in so doing, it provides "factual data for the formulation of public policies to deal with these problems" (NAS, 2013, p. 38).

Self-reports of SWB are likely to add useful information in instances where medical interventions have a desired outcome that is something other than merely an increase in life expectancy, where reflections of successful treatment and support extend beyond signs and symptoms and into domains such as functioning and social integration, and where parties other than the patients are affected by treatment and symptoms (care givers, family members, and others). See the article by Richard Frank article in this volume.

Future directions

Thinking in terms of a harmonized approach for national statistics offices to follow, the OECD *Guidelines* mark an important step forward in the measurement

[13] Dolan and Metcalfe (2011) surveyed people to ask whether government policy should seek to (1) improve happiness or (2) reduce misery, and there was more support for the second option.

of subjective well-being, but do not provide the 'final word' on the subject. Although some aspects of the measurement of subjective well-being – such as questions about overall satisfaction with life – are well understood, other potentially important measures currently draw on weaker evidence. It is expected that the evidence base will continue to develop rapidly over the next few years. In particular, to the extent that national statistical offices start regularly collecting and publishing data on SWB that researchers can exploit, many methodological questions are likely to be resolved, and an increasing body of knowledge will accumulate on the policy uses of these data.

National statistics offices face two issues in particular: (1) the need to pursue experimental techniques to push the state of the art forward; and (2) the need to collect high quality covariate data alongside SWB measures. Regarding the first, national statistics offices have long histories developing survey methods through systematic experiments and so are well positioned to contribute to the evolution of SWB measurement. While the OECD *Guidelines* were being drafted, the UK Office for National Statistics (ONS) was in the process of developing and collecting its first official measures of SWB. Typically, national statistical offices invest considerable methodological research upfront before collecting data for a new measure, but then implement collection in a homogenous way. In developing their measures of subjective well-being, the ONS deviated from this process significantly. Although the ONS did invest in methodological work before proceeding to measurement, rather than standardise on a single measure immediately, an experimental approach was taken by splitting the sample in their Household Opinion Survey and using this to test different questions, question order, and other methodological points. The experimental approach adopted by the ONS has had an important impact with respect to knowledge of the validity and reliability of subjective well-being measures and best practice with respect to question design.

On the second point, part of the experimentation process involves figuring out which subject matter domains (e.g. health, time-use, environment, city planning) benefit most from adding SWB content to existing surveys. A key advantage of many surveys carried out by government statistical agencies is that they generally collect higher quality information on potential covariates – such as income, labour force status, or education – than is possible in smaller unofficial surveys. Because the potential for insightful inferences to be drawn from SWB analyses depends not only on the quality of the SWB measure, but also on the quality of data on a range of other factors, surveys from national statistical agencies offer an opportunity to advance the field in a way that may not be possible elsewhere.

REFERENCES

BARTRAM, D., (2010). Economic migration and happiness: Comparing immigrants' and natives. Happiness gains from income. Social Indicators Research 103(1): 57−76.

BENJAMIN, D., KIMBALL, M., HEFFETZ, O., REES-JONES, A., (2013). Can Marginal Rates of Substitution be Inferred from Happiness Data? Evidence from Resiency Choices, NBER Working Paper 18927.

BOARINI, R., COMOLA, M., SMITH, C., MANCHIN, R., DE KEULENAER, F., (2012). What Makes for a Better Life? The determinants of subjective well-being in OECD countries: Evidence from the Gallup World Poll. STD/DOC(2012) 3, OECD.

BRADBURN, N., (1969). The Structure of Psychological Well-being. Chicago: Aldine.

CHRISTODOULOU, C., SCHNEIDER, S., STONE, A. A., (2013). Validation of a brief yesterday measure of hedonic well-being and daily activities: Comparison with the Day Reconstruction Method. Social Indicators Research.

CLARK, A. E., (2001). What really matters in a job? Hedonic measurement using quit data. Labour Economics 8: 223−242.

CLARK, A. E., OSWALD, A. J., (2002). A simple statistical method for measuring how life events affect happiness. International Journal of Epidemiology 31: 1139−1144.

CLARK, A. E., SENIK, C., (2011). Is Happiness Different from Flourishing? Cross-country Evidence from the ESS. Working Paper 2011-04, Paris, School of Economics.

CLARK, A. E., FRIJTERS, P., SHIELDS, M., (2008). Relative income, happiness, and utility: An explanation for the Easterlin paradox and other puzzles. Journal of Economic Literature 46(1): 95−144.

DEATON, A., (2010). Income, aging, health, and well-being around the world: Evidence from the Gallup World Poll. In Research Findings in the Economics of Aging, D. A. Wise (ed.), pp. 235−263. Chicago: University of Chicago Press.

DEATON, A., (2012). The financial crisis and the well-being of Americans. Oxford Economic Papers 64(1): 1−26.

DECI, E., RYAN, R., (2006). Hedonia, eudaimonia, and well-being: An introduction. Journal of Happiness Studies 9: 1−11.

DI TELLA, R., MACCULLOCH, R., OSWALD, A. J., (2001). Preferences over inflation and unemployment: Evidence from surveys of happiness. The American Economic Review 91(1): 335−341.

DI TELLA, R., MACCULLOCH, R., OSWALD, A. J., (2003). The Macroeconomics of Happiness. The Review of Economics and Statistics 85(4): 809−827.

DIENER, E., (2006). Guidelines for national indicators of subjective well-being and ill-being. Applied Research in Quality of Life 1(2): 151−157.

DIENER, E., (2011). The Validity of Life Satisfaction Measures. Unpublished manuscript, Department of Psychology, University of Illinois at Urbana-Champaign.

DIENER, E., CHAN, M., (2010). Happier People Live Longer: Subjective Well-Being Contributes to Health and Longevity. Unpublished manuscript, Department of Psychology, University of Illinois at Urbana-Champaign.

DIENER, E., PAVOT, W., (1993). Review of the satisfaction with life scale. Psychological Assessment 52(2): 164−172.

DIENER, E., HELLIWELL, J. F., KAHNEMAN, D. (eds.), (2010). International Differences in Well-Being. Oxford University Press.

DIENER, E., LUCAS, R. E., NAPA SCOLLON, C., (2006). Beyond the hedonic treadmill: Revising the adaptation theory of well-being. American Psychologist 61(4): 305−314.

DIENER, E., LUCAS, R. E., SCHIMMACK, U., HELLIWELL, J. (eds.), (2009). Well-Being for Public Policy. Oxford University Press.

DIENER, E., SUH, E. M., LUCAS, R. E., SMITH, H. L., (1999). Subjective well-being: Three decades of progress. Psychological Bulletin 125(2): 276−302.

DIENER, E., KAHNEMAN, D., TOV, W., ARORA, R., HARTER, J., (2009). Income's differential influence on judgments of life versus affective well-being. In Assessing Well-Being, E. Diener (ed.), pp. 233−246. Oxford, UK: Springer.

DOLAN, P., KAVETSOS, G., (2012). Happy Talk: Mode of Administration Effects on Subjective Well-Being (CEP discussion paper, no. 1159). London, UK: Centre for Economic Performance, London School of Economics and Political Science.

DOLAN, P., METCALFE, R., (2011). Comparing Measures of Subjective Well-Being and Views About the Role They Should Play in Policy. London, UK: Office for National Statistics.

DOLAN, P., PEASGOOD, T., (2006). Valuing non-market goods: Does subjective well-being offer a viable alternative to contingent valuation. Imperial College Working Paper. Imperial College, London.

DOLAN, P., WHITE, M., (2007). How can measures of subjective well-being be used to inform policy? Perspectives on Psychological Science. 2(1): 71−85.

DOLAN, P., LAYARD, R., METCALFE, R., (2011). Measuring Subjective Well-being for Public Policy. Office for National Statistics, UK.

DOLAN, P., PEASGOOD, T., WHITE, M., (2008). Do we really know what makes us happy? A review of the economic literature on the factors associated with subjective well-being. Journal of Economic Psychology 29: 94−122.

EASTERLIN, R., (1974). Does economic growth improve the human lot? Some empirical evidence. In Nations and Households in Economic Growth: Essays in Honour of Moses Abramovitz. P.A. David, and M.W. Reder (eds.), pp. 89−125. New York, Academic Press Inc.

FERRER-I-CARBONELL, A., FRIJTERS, P., (2004). How important is methodology for estimates of the determinants of happiness? The Economic Journal 114: 641−659.

FREY, B. S., STUTZER, A., (2002). What can economists learn from happiness research? Journal of Economic Literature 40(2): 402−435.

FREY, B. S., STUTZER, A., (2008). Stress that doesn't pay: The commuting paradox. Scandinavian Journal of Economics 110(2): 339−366.

FRIJTERS, P., (2000). Do individuals try to maximize general satisfaction? Journal of Economic Psychology 21: 281−304.

FUJIWARA, D., CAMPBELL, R., (2011, July). Valuation techniques for social cost-benefit analysis: Stated preference, revealed preference, and subjective well-being approaches, a discussion of the current issues. United Kingdom: HM Treasury and Department for Work and Pensions.

GERE, J., SCHIMMACK, U., (2011). A multi-occasion multi-rater model of affective dispositions and affective well-being. Journal of Happiness Studies 12(6): 931−945.

GRAHAM, C., (2011). The Pursuit of Happiness: An Economy of Well-Being. Washington, DC: The Brookings Institution Press.

HALPERN, D., (2010). The Hidden Wealth of Nations. Cambridge: Polity Press.

HELLIWELL, J. F., BARRINGTON-LEIGH, C. P., (2010). Measuring and Understanding Subjective Well-being. NBER Working Paper, No. 15887, National Bureau of Economic Research.

HELLIWELL, J. F., WANG, S., (2011). Trust and well-being. International Journal of Well-being 1(1): 42−78.

HELLIWELL, J. F., LAYARD, R., SACHS, J. D., (2013). The World happiness Report. The Earth Institute, Columbia University.

HUPPERT, F. A., SO, T. T. C., (2009). What Percentage of People in Europe are Flourishing and What Characterises Them? Well-Being Institute, University of Cambridge, mimeo prepared for the OECD/ISQOLS meeting on Measuring subjective well-being: An opportunity for NSOs? Florence, 23/24 July.

HUPPERT, F.A., MARKS, N., CLARK, A., SIEGRIST, J., STUTZER, A., VITTERSO, J., WAHRENDORF, W., (2009). Measuring well-being across Europe: Description of the ESS well-being module and preliminary findings. Social Indicators Research 91: 301−315.

KAHNEMAN, D., DEATON, A., (2010). High income improves evaluation of life but not emotional well-being. Proceedings of the National Academy of Sciences of the United States of America 107(38): 16489−16493.

KAHNEMAN, D., KRUEGER, A. B., (2006). Developments in the measurement of subjective well-being. Journal of Economic Perspectives 20(1): 3−24.

KAHNEMAN, D., DIENER, E., SCHWARZ, N., (1999). Well-being: The Foundations of Hedonic Psychology. Russell Sage Foundation, New York.

KAHNEMAN, D., KRUEGER, A. B., SCHKADE, D., SCHWARZ, N., STONE, A .A., (2004). A survey method for characterizing daily life experience: The Day Reconstruction Method (DRM). Science 306(5702): 1776−1780.

KAHNEMAN, D., KRUEGER, A. B., SCHKADE, D., SCHWARZ, N., STONE, A. A., (2006). Would you be happier if you were richer? A focusing illusion. Science 312(5782): 1908−1910.

KRUEGER, A.B., MUELLER, A., (2012). Time use, emotional well-being, and unemployment: Evidence from longitudinal data. American Economic Review 102(3): 594−599.

KRUEGER, A. B., SCHKADE, D., (2008). The reliability of subjective well-being measures. Journal of Public Economics 92(8−9): 1833−1845.

KRUEGER, A. B., STONE, A. A., (2008). Assessment of pain: A community-based diary survey in the USA. Lancet 371(9623): 1519−1525.

KRUEGER, A.B., KAHNEMAN, D., SCHKADE, D., SCHWARZ, N., STONE, A. A., (2009). National time accounting: The currency of life. In Measuring the Subjective Well-Being of Nations: National Accounts of Time Use and Well-Being, A. B. Krueger (ed.), pp. 9−86. Chicago: Chicago University Press.

LUCAS, R., (2007). "Long-Term Disability Is Associated With Lasting Changes in Subjective Well-Being: Evidence From Two nationally Representative Longitudinal Studies." Journal of Personality and Social Psychology, Vol. 92, (4), pp. 717−730.

LUCAS, R. A., CLARK, Y., GEORGELLIS, E., DIENER, (2004). "Unemployment alters the set point for life satisfaction", Psychological Science, 15, pp. 8−13.

NATIONAL RESEARCH COUNCIL, (2013). Subjective Well-Being: Measuring Happiness, Suffering, and Other Dimensions of Experience. Panel on Measuring Subjective Well-Being in a Policy-Relevant Framework. A.A. Stone and C. Mackie, Editors. Committee on National Statistics, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

NEW ECONOMICS FOUNDATION, (2009). National Accounts of Well-being.

OECD, (2011). How's Life? Measuring Well-Being. Paris, OECD Publishing.

OECD, (2013). OECD Guidelines on Measuring Subjective Well-being. Paris: OECD. Available: http://www.oecd.org/statistics/Guidelines%20on%20Measuring%20Subjective%20Well-being.pdf.

ONS, (2011). Initial Investigation into Subjective Well-being from the Opinions Survey.

OSWALD, A. J., WU, S., (2009). Well-Being Across America: Evidence from a Random Sample of One Million U.S. Citizens. Unpublished manuscript, University of Warwick, UK. Presented at the IZA Prize Conference, October 22, Washington, DC.

PAVOT, W., DIENER, E., COLVIN, C.R., SANDVIK, E., (1991). Further validation of the satisfaction with life scale: Evidence for the cross-method convergence of well-being measures. Journal of Personality Assessment 57(1): 149−161.

RAYO, L., BECKER, G. S., (2007). Evolutionary efficiency and happiness. Journal of Political Economy 115(2): 302−337.

ROBINSON, M. D., CLORE, G. L., (2002). Belief and feeling: Evidence for an accessibility model of emotional self-report. Psychological Bulletin 128(6): 934−960.

SAMPSON, R.J., GRAIF, C., (2009). Neighborhood social capital as differential social organization: Resident and leadership dimensions. American Behavioral Scientist 52(11): 1579−1605.

SCHULLER, T., WADSWORTH, M., BYNNER, J., GOLDSTEIN, H., (2012). The Measurement of Well-being: The Contribution of Longitudinal Studies. Report prepared for the Office for National Statistics. London. UK: Longview. Available: http://www.longviewuk.com/pages/documents/Longviewwellbeingreport.pdf.

SCHWARZ, N., STRACK, F., (1999). Reports of subjective well-being: Judgmental processes and their methodological implications. In Well-Being: The Foundations of Hedonic Psychology, D. Kahneman, E. Diener, and N. Schwarz (eds.), pp. 61−84. New York: Russell-Sage.

SMITH, T. W., (2005). Troubles in America: A Study of Negative Life Events Across Time and Sub-Groups (Russell Sage Foundation Working Paper Series).

STEVENSON, B., WOLFERS, J., (2013). Subjective well-being and income: Is there any evidence of satiation? American Economic Review 103(3): 598−604.

STEWART-BROWN, S., JANMOHAMED, K., (2008). Warwick-Edinburgh Mental Well-being Scale (WEMWBS). User Guide Version 1. Warwick and Edinburgh: University of Warwick and NHS Health Scotland.

STIGLITZ, J., SEN, A., FITOUSSI, J. P., (2009). Report by the Commission on the Measurement of Economic Performance and Social Progress. Available: http://www.stiglitz-sen-fitoussi.fr/documents/rapport_anglais.pdf.

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(63).

THOITS, P., (1983). Multiple identities and psychological well-being. American Sociological Review 48(2): 174−182.

THOMPSON, S., MARKS, N., (2008). Measuring Well-being in Policy: Issues and Applications. New Economics Foundation.

WHITE, M. P., DOLAN, P., (2009). Accounting for the richness of daily activities. Psychological Science 20(8): 1000−1008.

WILSON, T. D., GILBERT, D. T., (2006). Affective forecasting: Knowing what to want. Current Directions in Psychological Science 14(3): 131−134.

WILSON, T. D., WHEATLEY, T., MEYERS, J. M., GILBERT, D. T., AXSOM, D., (2000). Focalism: A source of durability bias in affective forecasting. Journal of Personality and Social Psychology 78(5): 821−836.