ISI 2011: SPECIAL TOPIC SESSION 26

MULTI-DIMENSIONAL MEASURES OF QUALITY OF LIFE

ABSTRACT

Making Happiness Count: four myths about measures of subjective well-being

Conal Smith

The notion that subjective perceptions are a fundamental component of quality of life is an old one. However, modern attempts to measure quality of life have traditionally emphasised those dimensions of life that could be measured objectively, and have particularly focused on command over resources. Recently this has begun to change, with an increasing interest in measures of subjective well-being from economists and sociologists. The 2009 *Report by the Commission on the Measurement of Economic Performance and Social Progress*, chaired by Nobel laureates Joseph Stiglitz and Amartya Sen, has placed measures of subjective well-being firmly at the centre of debates about how best to measure quality of life.

This paper addresses four common myths about measures of subjective well-being: that measures of subjective well-being are all about measuring happiness; that subjective well-being measures are not reliable or accurate; that respondents dislike being asked subjective questions in surveys; and that measures of subjective well-being have little immediate policy relevance. The first section of the paper outlines the emerging consensus around the core concepts captured by measures of subjective well-being and summarises the current state of empirical evidence on the validity, reliability, and accuracy of measures of subjective well-being.

The second section of the paper pools data from commercial surveys such as the Gallup World Poll and academic research initiatives such as the World Values Survey and the European Values Survey to analyse item non-response rates for a range of subjective and objective measures across multiple countries. This provides a large pool of quantitative evidence on respondent attitudes to subjective well-being questions across a different countries and cultures. The final section of the paper provides an overview of how measures of subjective well-being can be used to inform policy making.

MAKING HAPPINESS COUNT: FOUR MYTHS ABOUT MEASURES OF SUBJECTIVE WELL-BEING

Conal Smith

Introduction

The notion that subjective perceptions are a fundamental component of quality of life is an old one. Epicurus articulated such a position in ancient Greece, and Bentham made "the greatest happiness for the greatest number" the basis of utilitarian moral philosophy in the late 18th century. Since the 1870s, the conceptual underpinnings of modern economics have been based on a utilitarian framework that assumes people act to maximise their subjective preferences. More prosaically we implicitly acknowledge the significance of subjective perceptions whenever we ask a friend or relative "how are you" or respond to a doctor asking "tell me if this hurts".

Despite the centrality of subjective perceptions to both academic and day to day conceptions of quality of life, subjective well-being has traditionally been regarded as largely not measurable. Consequently, modern attempts to measure quality of life have usually emphasised those dimensions of life that could be measured objectively, and have particularly focused on command over resources. This is reflected in official statistics, where national statistical offices have, with a few notable exceptions, largely been reluctant to publish measures of subjective well-being.

Over the last two decades there has been increasing interest in measures of subjective well-being from economists and sociologists. This reflects a mounting weight of evidence that subjective well-being can be measured in a valid and reliable fashion. Following the academic literature on subjective well-being have come demands from policy makers and NGOs that measures of subjective well-being should be available to monitor progress and inform decision making. This has placed measures of subjective well-being firmly at the centre of debates about how best to measure quality of life. Perhaps the most high profile demand of this sort is from the 2009 *Report by the Commission on the Measurement of Economic Performance and Social Progress*, chaired by Nobel laureates Joseph Stiglitz and Amartya Sen. In particular, the Commission noted that:

> *Recent research has shown that it is possible to collect meaningful and reliable data on subjective well-being… National statistical offices should incorporate questions on subjective well-being in their standard surveys to capture people's life evaluations, hedonic experiences and life priorities.*[1]

In this environment there is considerable debate within national statistical offices over the place of measures of subjective well-being within official statistics. While a number of high profile initiatives relating to measures of subjective well-being have been launched over the past 18 months[2], many national statistical offices remain cautious. Some caution is warranted. In a tight fiscal environment resources for collecting new statistics are scarce. Furthermore, national statistical offices can reasonably claim they have a responsibility to only collect information that they are confident is both of high quality and for which there is a clear use. Nonetheless, much scepticism towards subjective measures is unfounded.

This paper addresses four common myths about measures of subjective well-being:

---

1 *Report by the Commission on the Measurement of Economic Performance and Social Progress*. Stiglitz, J. E., A. Sen and J.P. Fitoussi, 2009, p216.

2 ONS, INSEE, Eurostat

- measures of subjective well-being are all about measuring happiness;
- measures of subjective well-being are not valid or reliable;
- respondents dislike being asked subjective questions in surveys; and
- measures of subjective well-being have little immediate policy relevance.

The first two of these myths have been thoroughly addressed in the literature, and the relevant sections of this paper largely summarise what is known on the subject.

On the issue of respondent attitudes to subjective questions there is less in the literature, and the bulk of this section relies on the empirical analysis of several large cross-national datasets on subjective well-being. Using data from the Gallup World Poll, the World Values Survey, and the European Values Survey to analyse item non-response rates, the third section of the paper compares responses to subjective questions with those to more traditional questions such as income, education, and legal marital status.

The final section of the paper provides an overview of how measures of subjective well-being can be used to inform policy making.

1. Happiness and Subjective Well-being

In the media measures of subjective well-being are often conflated with or simply described as measures of "happiness". More disparagingly, the study of measures of subjective well-being is often labelled "happiology"[3]. The implication is that to be concerned with subjective well-being is to focus on trivial or fleeting emotions, and that subjective well-being is a fuzzy and imprecise topic.

The former view – that subjective well-being is trivial or fleeting – is built on the intuition that a full life involves more than just being happy. This line of argument too has a long tradition. In Homer‟s Odyssey, Odysseus‟ crew are detained in the „land of the Lotus eaters‟ when they consume the local food and fall into a happy, but apathetic state. It is clear from the poem that passive drugged happiness is one of the hazards that Odysseus must avoid in order to return home to Ithaca, rather than an outcome to be welcomed. Similarly, the framers of the US constitution were careful to identify the „pursuit of happiness‟ rather than its achievement as one of the fundamental rights that the constitution protects.

The second line of argument is that „happiness‟ is a vague and fuzzy topic that cannot be properly defined. Happiness, it is argued, means different things to different people. According to this line of argument it is not possible to put a precise definition around what constitutes happiness, and attempts to measure happiness are therefore fundamentally flawed.

Both arguments would have some weight if the measurement of subjective well-being primarily focused around some vaguely defined concept of "happiness". This, however, is not the case. In fact, the measurement of subjective well-being is about more than simply the measurement of "happiness", and there is an emerging consensus in the literature around the nature of the concepts to be measured (Sen, Stiglitz and Fitoussi, 2009, ONS, 2011). The framework used here identifies three broad concepts of well-being:

- life evaluation

[3] E.g. "A New Gauge to See What‟s Beyond Happiness", *New York Times*, May 16, 2011.

- affect
- eudaimonia (psychological "flourishing")

Life evaluations capture a reflective assessment of how one"s life is going. They are the result of a cognitive evaluation on the part of the subject rather than a description of current emotional state. A strength of measures of life evaluation is that they appear to capture the same underlying construct that people use when they decide that one course of action is preferable to another (Kahneman, 1999, Helliwell and Barrington-Leigh, 2010). It is for this reason that life evaluations are sometimes characterised as measures of "decision utility" (Kahneman and Krueger, 2006). However, this strength also comes with some disadvantages. In particular, life evaluations tend to be based on how we remember things, rather than how we experience them. Psychologists note that our memories of an experience tend to be characterised by the "peak/end rule" (Kahneman, 1999). The peak/end rule states that our evaluation of an experience tends to be dominated by the most intense (peak) emotion felt during the experience and the emotion felt at the end of the experience rather than on the average or integral of emotional intensity across the experience.

Many of the most commonly used measures of subjective well-being are evaluative measures. This reflects the fact that both academic economists and policy makers have a strong interest in the basis on which we make decisions, even if those decisions are based on how we remember things rather than how we experience them. Life evaluations also have the virtue that they tend to be relatively easy to measure in a household survey.

Affect is the term psychologists use to describe a person"s feelings. Measures of affect can be thought of as measures of particular feelings or emotional states, and is often measured with reference to a particular point in time. Such measures capture how we experience life rather than how we 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 hedonic dimensions: positive affect and negative affect. Positive affect captures positive emotions such as the experience of happiness, joy, and contentment. Negative affect, on the other hand, 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 that positive emotions are strongly correlated with each other), negative affect is more multi-dimensional. For example, it is possible at one given moment to feel anger but not fear or sadness.

The measurement of affect is more challenging than life evaluation. It is difficult to ask people to recall affective states in the past, since responses will be affected by recall biases such as the peak/end rule mentioned above. The gold standard for measuring affect is the experience sampling method (ESM), where an electronic device is used to prompt people to record their feelings and perhaps the activity they are undertaking at either random or fixed points over a period of time. While the ESM produces an accurate record of affect, it is also expensive to implement and intrusive for respondents. A more viable approach is the use of the day reconstruction method (DRM), in which respondents are questioned about events from a time use diary recorded on the previous day. Research has shown that the DRM produces results comparable with ESM, but with a respondent burden acceptable for including questions on affect in national time use surveys (Kahneman, Krueger, Schkade, Schwarz and Stone, 2004). At the cost of slightly less detail, it is also possible to obtain meaningful responses to questions in a standard household survey as to whether a person experienced particular affective states on the previous day.

Eudaimonic well-being comprises a range of different mental attributes and functionings that are thought to constitute mental "flourishing" (Huppert et al, 2009, NEF, 2009, Clark and Senik, 2011). This includes a sense of meaning or purpose in life, as well as feelings of agency and locus of control. While there is now a general consensus on the distinction between life evaluations and measures of 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 whether the term is used to cover a range of different sub-concepts. It is, however, clear that eudaimonic measures of well-being capture important aspects of our subjective perceptions about well-being not covered by life evaluations or affect. For example, having children has a negligible (or even mild 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, Krueger, Schkade, Schwarz and Stone, 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. However, people with children report much higher average levels of meaning or purpose in their lives (NEF, 2009).

Life evaluation, positive and negative affect, and eudaimonic well-being are all conceptually distinct. Are they empirically distinct? Table 1 below gives the correlations between measures of life evaluation (life satisfaction), positive affect, negative affect, and eudaimonic well-being (purpose) derived from the Gallup World Poll[4]. 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 about half as strong with both measures of affect, and half that with purpose. While all the coefficients show the expected sign and all are significant at the 0.1% level, for none of the measures is the correlation near 1 (See also Clark and Senik, 2011).

Table 1. Correlation Coefficients for Purpose, Life Satisfaction, Positive Affect, and Negative Affect, Gallup World Poll, 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

Feeling “happy” is an example of positive affect. However, adequately measuring subjective well-being requires separately measuring negative affect, life evaluations, and eudaimonic well-being also. As table 1 demonstrates, although these concepts are related to one another they are distinct. In a sense, criticisms of “happiology” are justified in that well-being is more than just happiness. However, the measurement of subjective well-being is not just the measurement of happiness.

2. Validity and Reliability

Quality is a priority for official statistics. For something to be included as part of official statistics, there must be both a clear conceptual grasp of what is to be measured, and also the ability to produce a measure of sufficient quality. The fact that there is a clear conceptual framework around a concept that we would like to measure does not, in and of itself, mean that a valid and reliable measure can be produced.

[4] 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.

This challenge is particularly acute for measures of subjective well-being. If the measure is subjective, how can we know objectively whether it is a good measure of the underlying concept?

At this point some precision is required about what is meant by a subjective measure. There are two senses in which we can talk about a subjective measure of something. A subjective measure can refer to either the measure itself, or to the concept being measured. In the first sense, it is possible to have a subjective measure of an objective concept. Consider, for example, the question "who do you think is older, John or Marama"? The measure is subjective in that it seeks a person‟s opinion, but the subject being measured (John and Marama‟s relative ages) is something that can be objectively verified. In particular, by checking the dates of birth for both John and Marama, different people can have equivalent access to objective information about the concept being measured. It is thus relatively trivial to test the validity of this sort of subjective measure simply by comparing the subjective responses to independent objective measures of the same outcome.

When the concept itself is subjective, however, things become a little more complicated. In the case of the question "how much do you like the colour blue", the concept being measured is itself subjective. There is no way for a person other than the respondent to have equivalent access to the concept being measured (the respondent‟s preference for blueness). This makes testing the validity of such measures much more challenging than in the first instance.

Measures of subjective well-being are subjective in this second sense, and this means we cannot simply compare measures of subjective well-being with clear objective measures of the same concept in order to reassure us of their validity. However, this does not mean that we cannot meaningfully analyse the validity of measures of subjective well-being at all. There is an extensive psychological literature on establishing the validity of subjective measures and this proposes at least three types of validity that a good subjective measure should demonstrate:

- Face validity (is the measure plausible?)
- Convergent validity (does the measure correlate well with other proxy measures for the same underlying concept?)
- Construct validity (does the measure perform in the way theory would suggest?)

The face validity of measures of subjective well-being is relatively straight forward to establish. The standard questions used have a clear intuitive relationship to the concept being measured. It is not a great stretch, for example, to suggest that asking a person whether they experienced sadness during the previous day is a plausible way to find out whether they felt sad during that day. However, there are a number of additional pieces of evidence that suggest that respondents find questions on subjective well-being easy to understand. Measures of subjective well-being have low item-specific non-response rates (Rässler and Riphahn, 2006), suggesting that respondent‟s do not find these types of question difficult to answer. This is supported by evidence that the time to reply is also low for measures of subjective well-being also (Diener and Tov, 2006).

Convergent validity involves examining whether a measure correlates well with other proxy measures for the same concept. Although measures of subjective well-being are focused around an inherently subjective concept, there is a range of information that we can use as proxy measures for people‟s subjective states. We can look at ratings, either by the respondent themselves over time or from other people. Similarly, we can observe the behaviour of the respondent to see if it is consistent with their reported subjective state. Finally, we can use biophysical measures related to emotion state. All of these

approaches have been applied to measures of subjective well-being and provide strong support for convergent validity.

- Respondents tend to report similar levels of subjective well-being when asked the same question at several different times. Test-retest results for subjective well-being measures yield correlations of between 0.6 and 0.7 between tests on the same day (Krueger and Schkade, 2007). Multiple item measures of subjective well-being do better than single questions, with test-retest scores close to 0.78 for time periods measured in weeks. These are lower than for some objective measures of economic variables, such as income (0.9), but not dissimilar from other more complex economic variables such as expenditure (0.6) measured over similar time periods (Carinna, G., D. Evans, F. Ravindal, and K. Xua, 2009).

- Ratings of a person‟s subjective well-being from friends and family have been shown to correlate well with self ratings of life satisfaction (Frey and Stutzer, 2002) as have, more surprisingly, ratings by the interviewer (Scheider and Schimmack, 2009). Similarly, for momentary affect, strangers shown a video or pictures of the respondent are able to accurately identify the subject‟s emotion (Diner, Suh, Lucas, and Smith, 1999). This latter finding applies to people from fundamentally different cultures to the respondent as well as people from the respondent‟s own culture.

- Subjective assessments of well-being are also reflected in behaviour. People who rate themselves as happy smile more. This applies particularly to so-called “Duchenne” or “unfakeable” smiles where the skin around the corners of the eye crinkles through a largely involuntary reflex (Frey and Stutzer, 2002). There is also good evidence that people act in ways that are consistent with what they say about their subjective well-being. That is, people avoid behaviour that they associate with a low level of subjective well-being (Frijters, 2000). For example, self-reports of job satisfaction have been shown to be a strong predictor of people quitting a job, even after controlling for wages, hours worked and other individual and job-specific factors.

- There have been a number of studies looking at the correlation between various bio-physical markers and subjective well-being. Measures of subjective well-being have been shown to be correlated with left/right brain activity (Urry et al., 2004), and with levels of the stress hormone cortisol in the bloodstream (Diener and Tov, 2006). People reporting high levels of subjective well-being recover more quickly from colds and minor injuries (Kahneman and Krueger, 2006).

Construct validity for measures of subjective well-being has been extensively researched. There is evidence that subjective well-being measures predict risk of suicide, sociability, extroversion, quality of sleep, and happiness of close relatives. Economists, driven in part by the desire to understand how well such measures function as a potential measure of utility, have looked in depth at the drivers of subjective well-being. It is clear that subjective well-being changes in the expected way in response to changes in circumstances such as changes in income, partnership status, health, or employment status (Dolan, Peasgood, and White, 2008). Further, it is clear that these changes are neither trivial in magnitude, nor transient. Studies have shown that change in income, becoming unemployed, and becoming disabled have a long lasting impact on subjective well-being (e.g. Lucas, 2007).

Table 2 below provides a summary of the evidence for the validity of subjective well-being outlined above.

Table 2. Evidence on the validity of measures of subjective well-being

Type of Evidence	**Sources**

Face Validity • Item-specific non-response rates • Time to reply	 Rässler and Riphahn, 2006 Diener and Tov, 2006
Convergent Validity • Self-ratings over time • Ratings by friends and family • Ratings from strangers • Frequency of smiling • Changes in behaviour • Biophysical measures	 Krueger and Schkade, 2007 Frey and Stutzer, 2002 Scheider and Schimmack, 2009, Diner, Suh, Lucas, and Smith, 1999 Frey and Stutzer, 2002, Kahneman and Krueger, 2006 Frijters, 2000 Urry et al., 2004, Diener and Tov, 2006
Construct Validity	Dolan, Peasgood, and White, 2008, Lucas, 2007, Helliwell, 2010

3. Respondent attitudes

National statistical offices are rightly cautious about alienating respondents. A strong rapport with respondents is of crucial importance both ethically and also pragmatically. The ethical issue reflects national statistical offices‟ obligation to use their legal authority in a responsible manner. The pragmatic concern is that, whether or not citizens are legally required to provide information to their national statistical office, voluntary responses are likely to be much better in terms of quality and cost than those obtained via compulsion. Moreover, if respondents see their national statistical office violating their trust in one area, this may carry over to their willingness to respond on other, unrelated topics. For this reason, a key concern with respect to producing official measures of subjective well-being is the acceptability of such measures to respondents.

Acceptability of this sort is less of an issue to academic researchers, who lack the legal powers of a national statistical office. Thus, while there is a large literature on the validity and conceptual basis for measures of subjective well-being, there is much less on the acceptability of such questions to respondents. What evidence there is generally supports the view that respondents do not object to answering such questions (Rässler and Riphahn, 2006). Nonetheless, it might be the case that asking such questions is less acceptable in some countries than others, or that the results cited previously are a product of the (relatively small) groups studied.

This section of the paper attempts to address the issue of the acceptability of questions on subjective well-being to respondents using three large cross-country datasets. The basic methodology is to compare both item-specific non-response rates (all reasons combined) and, more specifically, non-response due to a refusal to answer the question (as opposed to a “don‟t know” response), across countries and for a range of different measures. If people find questions on subjective well-being objectionable or too intrusive, we would expect to find relatively high item-specific non-response rates and, in particular, rates of refusal to answer. The paper examines both whether there are consistent patterns of non-response across countries, and how non-response rates for subjective well-being questions compare to other, more tradition questions included in household surveys.

The three datasets used here are sub-sets of the World Values Survey, the Gallup World Poll, and the European Values Survey. In the case of the Gallup World Poll and the World Values Survey countries have been dropped so the dataset includes only OECD countries along with Brazil, Russia, China, India, Indonesia, and South Africa. After this adjustment, the data from the World Values Survey covers 32 nations and involves 5 waves collected between 1981 and 2008. The total size of the dataset is 128,097 observations. Although the full Gallup World Poll dataset covers a total of 135 countries, the dataset used here covers a total of 40 countries. Five waves of data are used (2006-2010), comprising a total of 204,432

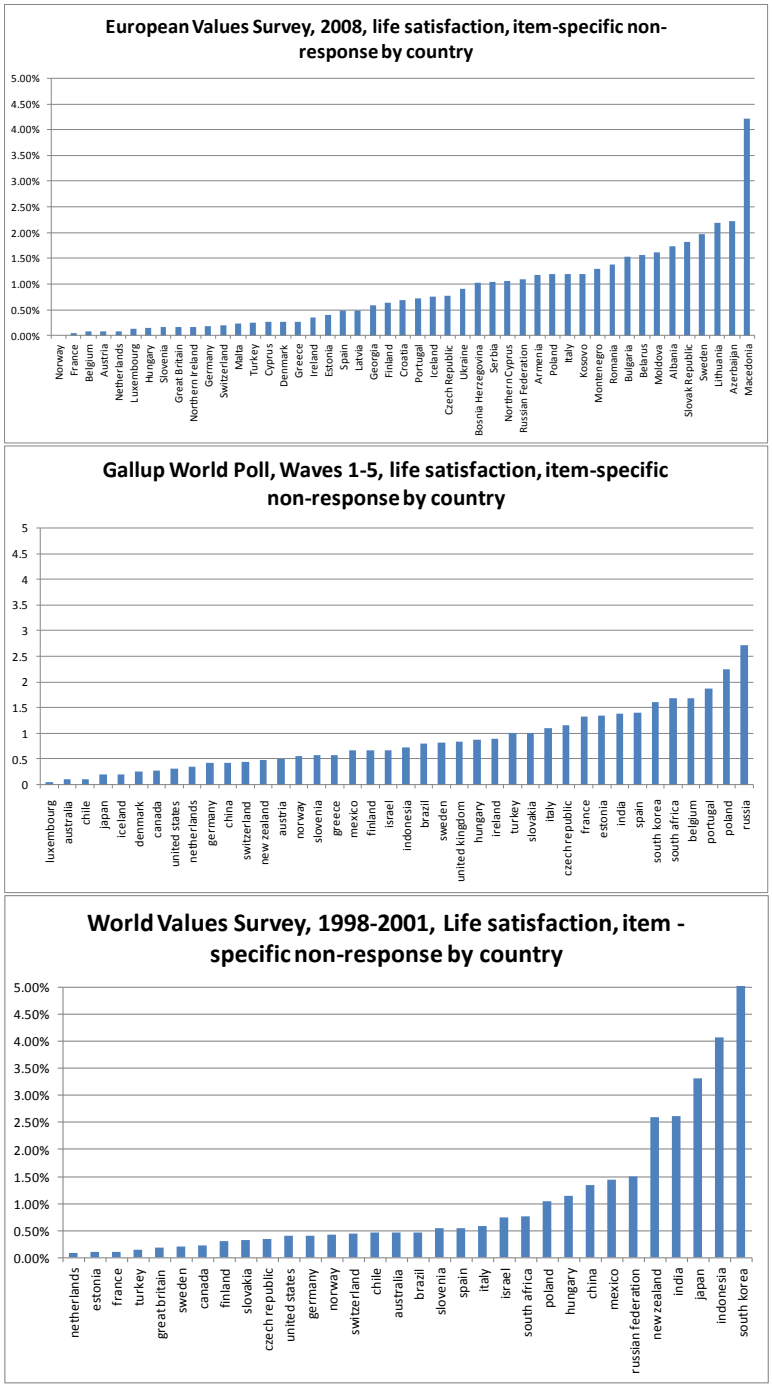
observations. The European Values Survey dataset is smaller than either of the above two surveys, covering 48 European countries but only one wave of data. Only wave 4 (2008) was used in this analysis, with a total sample size of 67,786 observations. For the Gallup World Poll and the European Values Study, the mode was largely computer assisted telephone interviewing. The World Values Study included both computer assisted telephone interviews and mail-out self-complete surveys in some countries.

Figure 1, below, shows the item-specific non-response rate (all reasons) for life satisfaction for all three surveys. From a quick glance at the figure it is evident that, with one exception, non-response rates are relatively low in each survey, with the majority of countries in each survey having less than 1% of respondents not responding. Indonesia (in the World Values Survey) and Macedonia (in the European Values Survey) have rates just over 4%, but are at the extreme tail of the distribution. South Korea has an item-specific non-response rate of 22.71% in the World Values Survey, but is a clear outlier[5]. Interestingly, neither Indonesia nor South Korea stand out in the Gallup World Poll with rates of just 0.72% and 1.6% respectively.[6]

[5] The World Values Survey observations on South Korea have an anomalously high non-response rate across a number of different items suggesting that the survey was not well-implemented in korea.

[6] No equivalent figure can be given for Macedonia, which is included in the European Values Survey but not in either of the other two datasets used here.

Figure 1. Total item-specific non-response rates for life satisfaction by country, European Values Survey, Gallup World Poll, and World Values Survey

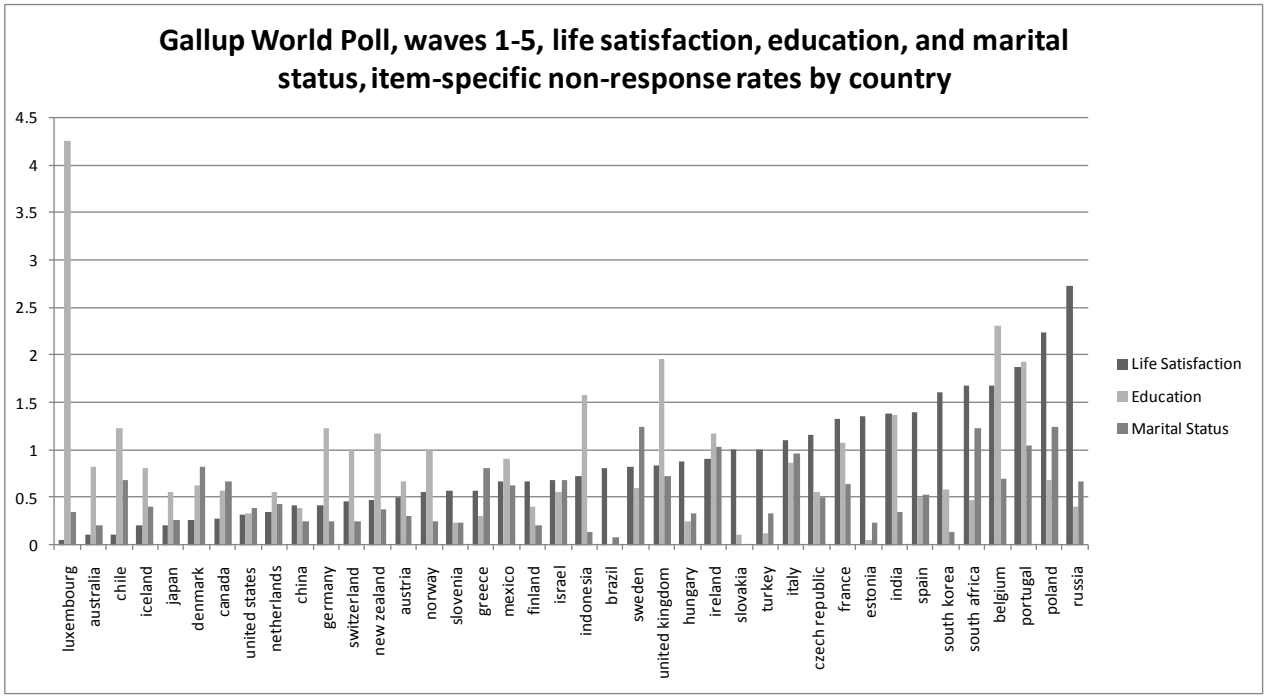


The three surveys have similar average levels of item-non response: 0.86% for the Gallup World Poll, 0.85% for the European Values Survey and 1.56% for the World Values Survey (0.88% if South Korea is dropped). However, although the average level of non-response to the life satisfaction question is similar in each survey, there is only a mild level of correlation between the levels of non-response on a country by

country basis[7]. Looking at the graphs there is a suggestion that northern European countries tend to have relatively low non-response rates and eastern European countries relatively high non-response rates, but this association is weak.

Does this (weak) association between countries and non-response to questions on life satisfaction support the view that there is a general dislike among respondents in some countries to answering questions on subjective well-being? Before it is possible to answer this question it is necessary to look at responses to other questions. Non response to questions on life satisfaction could represent an unwillingness to answer questions on life satisfaction, or an unwillingness to answer any sort of questions at all. Figure 2 below compares the item-specific non-response rates for life satisfaction with equivalent rates for marital status and education.

Figure 2. Total item-specific non-response rates for life satisfaction, education, and marital status by country, Gallup World Poll

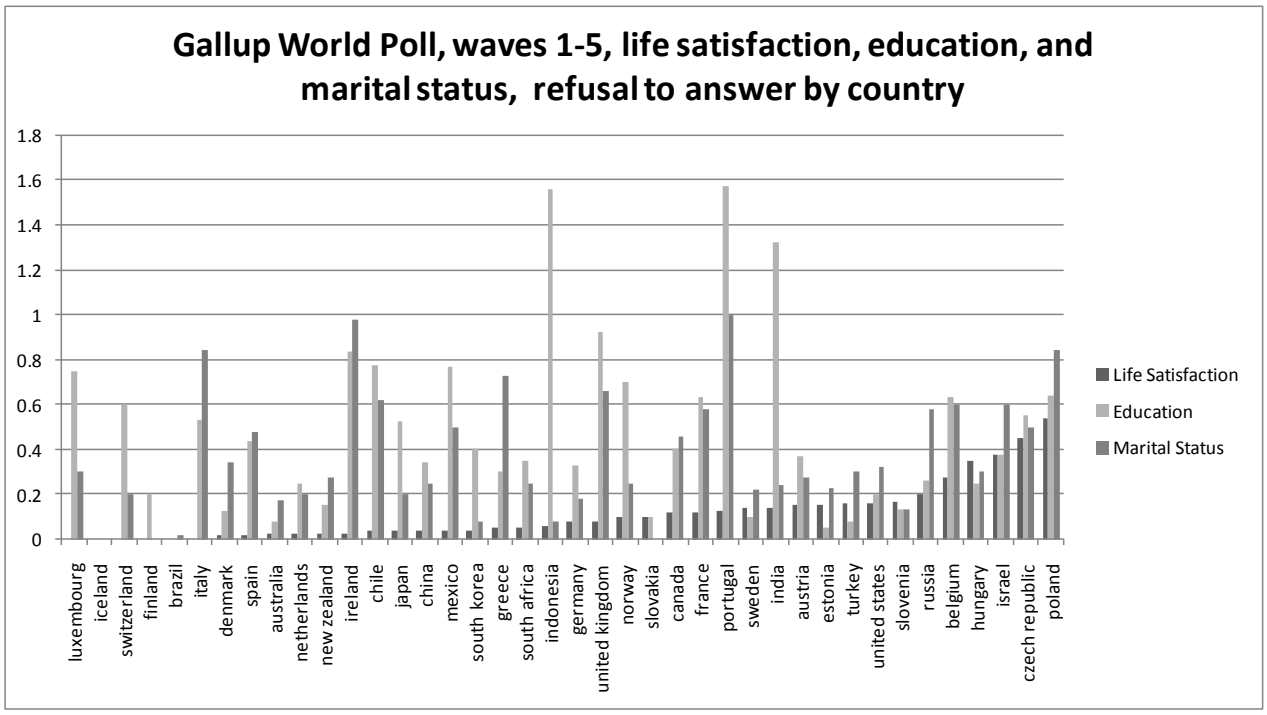


Two points are evident from figure 2. First, life satisfaction has generally a similar level of non-response to marital status and education. In fact, the average non-response rate for education and life satisfaction across countries is very close at 0.86 for life satisfaction and 0.85 for education. Non-response for marital status is a little lower at an average of 0.51. The standard deviations for the three variables are also similar (0.72, 0.90, and 0.65 respectively for life satisfaction, education, and marital status), suggesting that there is nothing particularly unusual in the level of variation in non-response to life satisfaction.

[7] Correlation coefficients range from 0.37 between the Gallup World Poll and the European Values Survey, to 0.02 between the Gallup World Poll and the World Values Survey (omitting South Korea). The correlation between the European Values Survey and the World Values Survey results is 0.19. Of these, only the relationship between the European Values Survey and the Gallup World Poll is significant at the 10% level (and even this fails at the 5% level). Including South Korea, the relationship between the Gallup World Poll and the World Values Survey is still insignificant.

Even if non-response rates are broadly similar for life satisfaction and other questions, it might still be the case that respondents are more willing to answer questions on subjects like education and marital status, but also find these questions more difficult to answer. In other words, similar levels of total non-response might conflate dislike of answering subjective questions and difficulty in accurately answering more objective questions. It is possible to cast some light on whether this is the case by distinguishing between non-response due to a refusal to answer, and non-response due to the respondent indicating that they "don't know" the answer. Figure 3 presents essentially the same information as figure 2, but based on refusal to answer rather than total non-response.

Figure 3. Rate of refusal to answer for life satisfaction, education, and marital status by country, Gallup World Poll



It is immediately evident from figure 3 that, when the focus is narrowed from total non-response to refusal to answer the relative performance of life satisfaction actually improves. In fact, life satisfaction has a lower refusal to answer rate than education in all countries in the dataset except Slovenia, Slovakia, and Hungary, and a lower rate than marital status in all countries except Estonia, Slovenia, Hungary, Sweden, and Turkey. These relationships are not unique to the Gallup World Poll either. Repeating the same comparison with the European Values Survey produces qualitatively similar results[8] in that refusal to answer is, on average, higher for education and marital status than life satisfaction. However, the countries that stood out in the Gallup World Poll as having higher rates of refusal to answer for life satisfaction than for other questions do not do so in the European Values Survey.

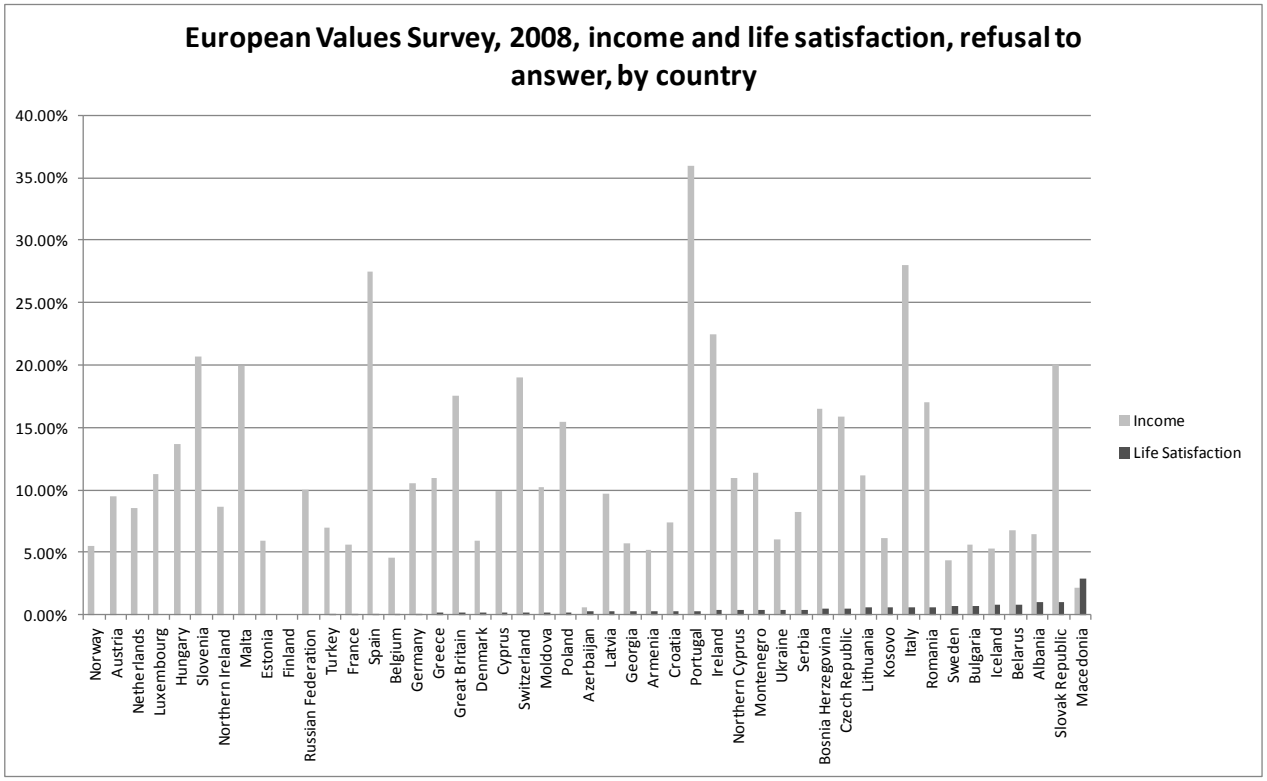
The overall picture, then, is that both non-response rates and, more specifically, refusal to answer is not particularly high for life satisfaction either in absolute terms or compared to more standard questions such as education or marital status. Similarly, there is little consistency in the pattern of variation either

[8] The version of the World Values Dataset used here does not permit breaking non-response down into "don't know" vs "refused to answer".

across countries or across questions, suggesting that the variation we see in non-response rates is random variation rather than consistent differences in national temperament.

It is, however, worth considering one more example in order to demonstrate the difference between a question that respondents are relatively happy to answer compared to one that they are not. Figure 4 compares the total non-response rate and the rate of respondents refusing to answer for income and life satisfaction.

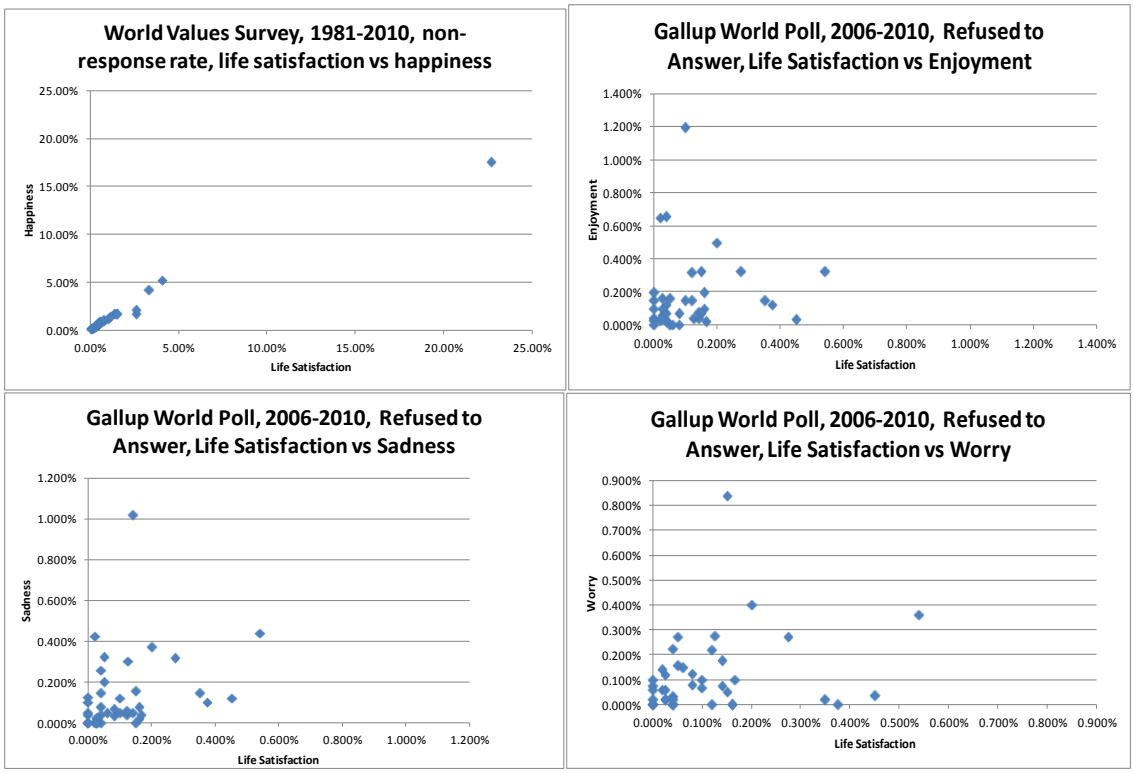
Figure 4. Rate of refusal to answer for life satisfaction and income by country, European Values Survey



As should be relatively clear from figure 4, it is difficult to meaningfully compare non-response to income and life satisfaction. Rates of refusal to answer are, for the majority of countries, between 10 and 100 times higher for income than for life satisfaction. Analysis of item-specific non-response (all reasons) using the World Values Survey produces similar results. Put simply, income (which is collected regularly and successfully as part of official statistics across the world) is a vastly more sensitive topic for respondents than subjective well-being.

All of the analysis so far has focused on life satisfaction. It is, however, possible that respondents are happy to answer questions on life evaluations, of which life satisfaction is an example, but would be more reluctant to answer questions about affect (current feelings). The latter might be deemed more personally intrusive, or it might simply be more difficult for respondents to understand how data on their current feelings could be a subject of valid interest for a household survey. Figure 5 below shows scatterplots of refusal to answer for life satisfaction and enjoyment, worry, and sadness from the Gallup World Poll and for non-response for life satisfaction and happiness from the World Values Survey.

Figure 5. Non-response rates and rates of refusal to answer for life evaluations and affect measures by country, World Values Survey and Gallup World Poll



It is evident from figure 5 that the happiness and life satisfaction questions generate essentially the same levels of non response from respondents to the World Values Survey. However, this is unsurprising given that the wording of the happiness question in the World Values Survey puts it in an intermediate position between a question clearly focusing on life evaluation (such as life satisfaction) and a pure affect question. The affect questions in the Gallup World Poll, however, are very clearly focused on current feelings. These do generate a different set of refusals to respond when compared to life satisfaction[9], and are, on average, slightly higher. Despite this, the overall magnitude of non-response is similar to that for life satisfaction, suggesting that most of the conclusions relating to the acceptability of life satisfaction apply also to measures of affect.[10]

All the quantitative evidence available from three separate surveys spanning 30 years and 61 countries supports the view that respondents do not object to answering questions on subjective well-being. Equally, there is no evidence to support the view that there is a consistent pattern of differences in attitudes across countries with people from some countries consistently more unwilling to answer such questions than those in others. Indeed, subjective questions perform very similarly to standard demographic questions such as marital status or education and perform vastly better than income. Questions on affect produce generally similar results in terms of levels of non-response to questions on life satisfaction.

[9] None of the pairwise correlation coefficients are significant and all are very small in absolute terms.

[10] For reasons of space, and due to the limited range of eudaimonic questions contained in the datasets used here, no analysis has been done on respondent attitudes to questions on eudaimonic well-being.

4. Policy Uses

The demand for statistics is a derived demand. Even if subjective well-being can be validly measured, and respondents are happy to be asked about it, this still leaves the question of whether the data is actually useful. If there is no practical policy use for the data, then there is no point in collecting it. It is not unreasonable, therefore, for national statistical offices to wish to establish how subjective measures could be used before deciding whether or not to commit to collecting them.

There is a growing literature on the policy relevance of measures of subjective well-being. Some authors (e.g. Kenny, 2011) question whether subjective well-being is a meaningful policy goal given that many of the determinants of subjective well-being lie outside the control of government and that average levels of subjective well-being do not change rapidly over time. However, neither objection is particularly compelling. Many (arguably the majority) of the determinants of economic growth, migration rates, population health, and crime rates lie outside the control of government, yet these areas are widely considered at the core of public policy and drive much of the demand for official statistics.

The belief that measures of subjective well-being are only useful for policy if they change rapidly over time appears to be based on the implicit assumption is that subjective well-being measures should be used to inform policy in a manner analogous to GDP. That is, policy makers should monitor changes in the level of subjective well-being and intervene in order to make sure that subjective well-being continually increases. However, such views are based on a fundamental misconception about why measures of subjective well-being are important for policy making.

Measures of subjective well-being are important because they provide relevant information that other, more traditional, measures cannot. In particular, measures of subjective well-being can:

- *complement existing progress measures* at an aggregate national level
- enable us to identify empirically *what matters for well-being* at the level of the individual, and quantify the importance of different outcomes
- provide the empirical foundation for better *cost-benefit analysis*, particularly where non-market outcomes are involved
- assist in *understanding human behaviour and decision making*.

Complement existing progress measures

Measures of subjective well-being provide an alternative and complementary measure of overall progress to more conventional measures that is firmly grounded in things that actually matter to people. Because it provides an overall picture, subjective well-being can help identify situations where more traditional indicators are missing something important by highlighting that different measures are moving in different ways. For example, data from the Gallup World Poll demonstrates a mismatch between improving economic and health outcomes on one hand, and falling levels of life satisfaction on the other, in Egypt and Tunisia in the years prior to the democratic revolutions of 2010 and 2011 (Gallup, 2011).

What matters for well-being

Subjective well-being measures can be used to test empirically which objective conditions are significant components of individual well-being. With appropriate analysis this may allow the estimation of the relative importance of different factors and how this impact may differ across different population

groups. Without using subjective well-being analysts are left with theoretical accounts and expert a-priori judgements about what matters and what is most important to individuals.

In general, the evidence from subjective well-being supports intuititive views as to what matters to people – health, income, not being unemployed, and social contact are all important (Dolan, Peasgood, and White, 2008). However, it can also provide some more surprising results: procedural issues have an intrinsic contribution to well-being not just an instrumental one. For example, both feelings of confidence in government and public institutions (Helliwell, 2008) and the level of direct democratic engagement through referenda and other means (Frey and Stutzer, 2000) are associated with high levels of subjective well-being.

Cost-benefit analysis

The use of measures of subjective well-being allows for better treatment of non-monetary outcomes in cost-benefit analysis (e.g. Clarke and Oswald, 2002, Dolan and Metcalf, 2008). Existing ways of obtaining values of non-monetary outcomes for cost-benefit analysis – such as willingness to pay or the use of shadow prices – are expensive to collect and known to produce results which are neither always intuitively plausible, nor internally consistent, and can be heavily affected by strategic decisions on the part of those surveyed (Dolan and Metcalfe, 2008). Measures of subjective well-being are cheaper to collect, produce relatively consistent results, and are mostly immune to strategic manipulation by respondents. For example, by looking at the marginal impact on subjective well-being of say, an improvement in health status as opposed to moving from employment to unemployment, it is possible to estimate the effects of a given output on well-being. Combining this information with figures for the cost of purchasing each output, it is possible to estimate the relative benefits of a dollar spent on the health intervention as opposed to the active labour market programme. At a programme level, subjective well-being measures can be included as outcome measures in experimental or quasi-experimental evaluations of policy programmes, providing stronger causal evidence on the impact of a policy on well-being.

Understanding human behaviour and decision making

Analysis of subjective well-being measures can shed light on some of the fundamental assumptions about human behaviour. For example, the axiom that efficiently functioning markets maximize well-being depends crucially on the extent to which people are able to make choices that will enhance their well-being. Subjective well-being measures can take debate beyond a simple theoretical debate about whether there are externalities present or whether people are fully rational, and enable analysis of what sorts of errors people actually make in forecasting their future affective states, and how significant errors in judgements of this sort are compared to other factors. This type of information is crucial to policy making, since public policy is largely focused on altering the behaviour of individuals. For example, retirement income policy is largely justified on the basis that individuals are not good judges of their own future well-being, or that their future well-being is given relatively little weight in decision making compared to current well-being.

5. Concluding Remarks

The objective of this paper has been to briefly cover some of the main concerns that national statistical offices might raise with respect to collecting official data on measures of subjective well-being. The issues dealt with here are not exhaustive, and address only some of the questions that might be raised about subjective data. Furthermore, much of this paper merely summarises findings from the existing literature.

What then, are the main points to take from this review? There are four. First, the measurement of subjective well-being should not be conflated with a naïve "happiology". There is a clear conceptual framework underlying the measurement of subjective well-being, and this framework covers a range of concepts as well as narrowly defined "happiness". There is good evidence to support the view that different concepts correspond well with the processes via which people make decisions and their experience of life as it unfolds.

Second, there is an extensive and solidly grounded literature suggesting that measures of subjective well-being are conceptually sound and methodologically valid. This is not to say that measures of subjective well-being are entirely unproblematic. However, the limitations of subjective measures are not so extreme that such measures should necessarily be considered beyond the scope of official statistics.

Third, it is clear that there is no evidence of significant respondent reluctance to answer questions on subjective well-being. In fact, the evidence strongly suggests that questions on subjective well-being are much less sensitive for respondents than some of the core concepts measured by official statistics such as income. This conclusion applies across all the countries covered by the datasets used, and is not dependent on the choice of dataset. When the size and scope of the datasets used to reach this conclusion is considered, along with the discrepancy in non response between income (which is regularly and successfully collected as part of official statistics) and measures of subjective well-being, it should be clear that the burden of proof sits squarely with those raising concerns about the willingness of respondents to answer questions on subjective well-being.

Finally, measures of subjective well-being have a range of different policy applications. The demand for measuring subjective well-being should not be viewed solely through the lens of "alternatives to GDP", but needs to take into account the crucial role that such measures can play in helping policy-makers evaluate the relative importance of fundamentally different outcomes. This is particularly important with respect to taking account of non-market outcomes in cost-benefit analysis.

The demand for measures of subjective well-being can be expected grow as both policy makers and the general public become more familiar with the use of such measures. How and when national statistical offices respond to this demand will vary from country to country. However, it is important that as national statistical offices develop their responses, that these are grounded in a solid understanding of the strengths and weakness of measures of subjective well-being rather than in a series of myths.

Bibliography

Bjornskov C. (2010), "How Comparable are the Gallup World Poll Life Satisfaction Data?", *Journal of Happiness Studies*, Vol. 11, pp 41-60.

Blanchflower D. G. and A. J. Oswald, (2004) "Money, Sex and Happiness: An Empirical Study", *Scandinavian Journal of Economics*, Vol. 106 (3), pp 393-415

Carinna, G., D. Evans, F. Ravindal, and K. Xua (2009), "Assessing the reliability of household expenditure data: Results of the World Health Survey, *Health Policy 91*, pp297–305

Clark, A. and A. Oswald (2002), "A simple statistical method for measuring how life events affect happiness", *International Journal of Epidemiology 31*, pp 1139-1144

Clark, A. and C. Senik (2010) "Will GDP growth increase happiness in developing countries?", *Paris School of Economics Working Paper No 2010 - 43*

Clark, A. and C. Senik (2011) "Is happiness different from flourishing? Cross-country evidence from the ESS", *Working Paper 2011-04*, Paris School of Economics

Cummins, R., R. Eckerslet, J. Pallant, J. Vugt, and R. Misajon, R (2003), "Developing a National Index of Subjective Wellbeing: the Australian Unity Wellbeing Index", *Social Indicators Research* 64, pp. 159-190

Cummins, T. and E. Gullone (2000), "Why we should not use 5-point Likert scales: the case for subjective quality of life measurement", *Proceedings of the Second International Conference on Quality of Life in Cities*, Singapore National University, pp74-93

Diener E. (1984), "Subjective Well-Being", *Psychological Bulletin*, No. 93, pag. 542-575

Diener E. and W. Tov (2006), "National Accounts of Well-being", *Encyclopedia of Quality of Life*, K. Land (eds)

Diener E., J. F. Helliwell and D. Kahneman (2010), *International Differences in Well-Being*, Oxford University Press, New York

Dolan P., T. Peasgood and M. White (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*, Vol. 29, pp 94-122.

Dolan, P. and Metcalfe, R (2008), *Comparing Willingness-To-Pay and Subjective Well-Being in the Context of Non-Market Goods*, CEP Discussion paper No 890, London School of Economics.

Easterlin, R. (1974) "Does Economic Growth Improve the Human Lot? Some Empirical Evidence", in David, P. A. and M. W. Reder, *Nations and Households in Economic Growth: Essays in Honour of Moses Abramovitz*, New York, Academic Press Inc, pp89-125.

Ferrer-i-Carbonell, A. and P. Frijters (2004), "How important is methodology for the estimates of the determinants of happiness?", *The Economic Journal*, Vol. 114, pp 641-659.

Fox, C. and Kahneman, D (1992) "Correlates, Causes and Heuristics in Surveys of Life Satisfaction", *Social Indicators Research* 27, 1992, Pp221 – 234

Frey, B. S. and A. Stutzer (2002), "What Can Economists Learn from Happiness Research?", *Journal of Economic Literature 40 (2)*, pp. 402-435

Frey, B. S. and A. Stutzer (2008), "Stress that Doesn't Pay: The Commuting Paradox", *Scandinavian Journal of Economics*, Vol. 110 (2), pp 339-366.

Frey, B. S. and A. Stutzer, (2000), "Happiness, Economy and Institutions", *The Economic Journal*, Vol 110. (466), pp 918-938.

Frijters, P. (2000), "Do individuals try to maximize general satisfaction?" *Journal of Economic Psychology 21*, pp 281-304

Helliwell, J. F. & C. P. Barrington-Leigh (2010), "Measuring and Understanding Subjective Well-being", *NBER Working Paper 15887*, National Bureau of Economic Research,

Helliwell, J. F. (2008) "Life Satisfaction and the Quality of Development", *NBER Working Paper 14507,* National Bureau of Economic Research.

Huppert, F., N. Marks, A. Clark, J. Siegrist, A. Stutzer, J. Vitterso and M. Wahrendorf (2009), "Measuring Well-being Across Europe: Description of the ESS Well-being Module and Preliminary findings", *Social Indicators Research,* pp301-315

Inglehart, R & H. Klingemann (2000), "Genes, Culture, Democracy, and Happiness." Ed Diener and E.M. Suh (eds), *Subjective Well-Being Across Cultures*, MIT Press, Cambridge MA.

INSEE (2010), "*Un measure de la qualité de vie",*XXX.

Kahneman D., E. Diener and N. Schwaarz (1999), *Well-being. The Foundations of Hedonic Psychology*, Russel Sage Foundation, New York.

Kahneman, D. and A. B. Krueger (2006), "Developments in the Measurement of Subjective Well-Being." *Journal of Economic Perspectives,* Vol. 20 (1), pp19-20.

Kahneman, D. and A. Deaton (2010), "High income improves life evaluation but not emotional well-being", Proceedings of the National Academy of Sciences, Vol. 107 (38), pp 16489-16493.

Kahneman, D., A. Krueger, D. Schkade, N. Schwarz, and A. Stone, (2004) "A Survey Method for Characterizing Daily Life Experience: The Day Reconstruction Method", *Science 3* Vol. 306. (5702), pp. 1776 - 1780

Kapteyn, A., J. P. Smith, and A. van Soest (2010), "Are Americans Really Less Happy With Their Incomes?", Paper prepared for the 31st General Conference of the International Association in Income and Wealth

Krueger, A. B. and D. A. Schkade (2007), *The Reliability of Subjective Well-being Measures,* NBER Working Paper 13027.

Larson, R. J. and B. L. Fredrickson (1999), "Measurement Issues in Emotion Research" in D. Kahneman, E. Diener, & N. Schwarz (eds), *Well-being. The Foundations of Hedonic Psychology*, Russel Sage Foundation, New York, pp40-60

Lucas, R. 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). pp717-730.

New Economics Foundation (2009), *National Accounts of Wellbeing,* January, pp. 55-56

Pudney, S. (2010), *An experimental analysis of the impact of survey design on measures and models of subjective wellbeing*, Institute for Social and Economic Research.

Rässler, S. & R. T. Riphahn (2006), "Survey item nonresponse and its treatment", *Allgemeines Statistisches Archi,v* 90, pp 217-232.

Sacks, W. D., B. Stevenson and J. Wolfers (2010), "Subjective Well-being, Income, Economic Development and Growth", NBER Working Paper 16441.

Shneider, L., and U. Schimmack (2009), "Self-Informant Agreement in Well-Being Ratings: A Meta-Analysis", *Social Indicators Research 94*, pp363–376

Stevenson, Wolfers, (2008) "Economic Growth and Subjective Wellbeing: Reassessing the Easterlin Paradox", *NBER Working Paper No. 14282*

Stiglitz J. E., Sen A., Fitoussi J.P. (2008), "Report by the Commission on the Measurement of Economic Performance and Social Progress".

Torangeau, R., Rasinski, K. A., and Bradburn, N. (1991), "Measuring Happiness in Surveys: A Test of The Subtraction Hypothesis", *Public Opinion Quarterly* 55, pp 255-266

Tourangeau, R, (1999), "Context Effects on Answers to Attitude Questions" in *Cognition and Survey Research*, Sirken, Monroe G. Ed Wiley, New York, Pp 111-131

Urry, H., J. Nitschke, I. Dolski, D. Jackson, K. Dalton, C. Mueller, M. Rosenkranz, C. Ryff, B. Singer, and R. Davidson (2004), "Making a Life Worth Living: Neural Correlates of Well-Being", *Psychological Science 15 (6)*, pp 367-372

Van Praag, B.M.S., P. Frijters and A. Ferrer-i-Carbonell (2003), "The anatomy of subjective well-being", *Journal of Economic Behaviour and Organisation*, 51, pp. 29-49

Winkelman, L. and R. Winkelman (1998), "Why Are The Unemployed So Unhappy? Evidence From Panel Data?", *Economica*, Vol. 65. pp 1- 15.

Winkelmann, R. (2004) "Subjective Well-Being and the Family. Results from an Ordered Probit Model with Multiple Random Effects." University of Zurich.