The Science of Happiness for Policymakers:

An Overview

Dr Dan Weijers
Philosophy Programme
Victoria University of Wellington
danweijers@gmail.com

Abstract

With several national governments and multinational organisations investigating new measures of progress and well-being to inform policymaking, some researchers have called for measures of happiness to be among those investigated. This paper provides a high-level tour of this debate, which includes important background information and discussion of some objections to the most promising measures of happiness. Following this, I argue that measures of happiness should play an important role in policymaking. However, I also note that several important obstacles need to be overcome before any measure of happiness could play such a role. The problems are manifold and will require considerable interdisciplinary work to overcome. Nevertheless, it is clear that the importance of happiness should make this work a priority.

Keywords: well-being, happiness, subjective well-being, public policy, science of happiness
1. Introduction

Traditional economic indicators of progress are widely seen to be insufficient as indicators of well-being (Michalos 2011; Stiglitz, Sen, & Fitoussi 2009). Even if the traditional economic measures of per capita production, income, and wealth took all relevant production, income, and wealth into account (which they don’t), they would still fail to capture the value of our relationships, health, and happiness—all of which are typically viewed as important for well-being. Many alternate measures of well-being have been developed over the years to address this problem, and governments have slowly incorporated some of them into the policymaking process at various stages. Over the last few years, the debate about whether happiness should be measured and used as an indicator of progress, and to inform policymaking, has intensified. This paper provides a high-level tour of this debate, includes an investigation into the severity of the problems of using measures of happiness for policymaking, and pays particular attention to measures of subjective well-being (predominantly survey questions about how happy or satisfied respondents are with their lives).

First, I discuss how we have arrived at the point where using measures of happiness is being seriously considered by policymakers. Then I explain the most promising methods for measuring happiness. Following this, I discuss several important criticisms of these measures, some of which are argued to be problems that need addressing before measures of happiness can usefully be employed by policymakers. The main criticisms addressed include: survey measures of happiness are too insensitive, we cannot know what measures of happiness are really measuring, and that the wrong kind of happiness is being measured. Finally, I provide recommendations for the role that suitably-improved measures of happiness could and should play in policymaking and what steps should be undertaken to suitably improve these measures.

This paper concludes that happiness should indeed be measured and used by governments and civil servants to inform policymaking. However, much complex interdisciplinary and international research will be required before measures of happiness can fruitfully play such a role.
1.1 Why Measure Happiness?

The limits of Gross Domestic Product (GDP) and Gross National Product (GNP) have been recognised by economists, politicians, and others for a long time (Michalos 2011). In particular, GDP and GNP have been criticised as measures of progress because of their myopic focus on production. These measures were never intended to be complete measures of progress, rather they were intended to reveal the rate at which the economy was growing or shrinking in a country (England 1998). However, the pursuit of economic growth has dominated the agendas of nearly all national governments and as a result GDP became the main indicator of progress almost by default. Some of the problems with GDP and GNP were emotively elucidated by Robert F. Kennedy in his famous speech at the University of Kansas in 1968:

“But even if we act to erase material poverty, there is another greater task; it is to confront the poverty of satisfaction—purpose and dignity—that afflicts us all. Too much and for too long, we seemed to have surrendered personal excellence and community values in the mere accumulation of material things. Our Gross National Product, now, is over $800 billion dollars a year, but that Gross National Product—if we judge the United States of America by that—that Gross National Product counts air pollution and cigarette advertising and ambulances to clear our highways of carnage. It counts special locks for our doors and the jails for the people who break them. It counts the destruction of the redwood and the loss of our natural wonder in chaotic sprawl. It counts napalm and counts nuclear warheads and armored cars for the police to fight the riots in our cities. It counts... the television programs, which glorify violence in order to sell toys to our children. Yet the gross national product does not allow for the health of our children, the quality of their education or the joy of their play. It does not include the beauty of our poetry or the strength of our marriages, the intelligence of our public debate or the integrity of our public officials. It measures neither our wit nor our courage, neither our wisdom nor our learning, neither our compassion
nor our devotion to our country; it measures everything, in short, except that which makes life worthwhile. And it can tell us everything about America except why we are proud that we are Americans.”

In response to these and other shortcomings of GDP and GNP, economists, national statisticians, and interested non-governmental organisations began to investigate and measure many other aspects related to progress. In addition to broadening and refining the existing range of economic measures, these initiatives also led to the collection of data related to individual well-being, quality of life, and even happiness. For many years now, academics from several disciplines and various countries, and even some civil servants, have been increasingly pushing for these new measures of well-being to play more important roles in policymaking. And, over the last few years, politicians have finally begun to listen. In 2008, French President Nicolas Sarkozy chartered the Commission on the Measurement of Economic Performance and Social Progress. The Commission, headed by Nobel-winning economist Joseph Stiglitz, advised that their report is...

... addressed, first of all, to political leaders. In this time of crises, when new political narratives are necessary to identify where our societies should go, the report advocates a shift of emphasis from a “production-oriented” measurement system to one focused on the well-being of current and future generations, i.e. toward broader measures of social progress. (Stiglitz, Sen, & Fitoussi 2009, p. 10)

The Commission mentions measures of subjective well-being briefly, but positively, encouraging national statistical offices to “incorporate questions to capture people’s life evaluations, hedonic experiences and priorities in their own surveys” because “[m]easures of subjective well-being provide key information about people’s quality...

---


of life.” (Stiglitz, Sen, & Fitoussi 2009, p. 58). The commission’s report seems to have had the required effect on British Prime Minister David Cameron, who gave a speech announcing the investigation of and subsequent measurement of well-being (including subjective well-being) by the British government on 25 November 2010:

[F]rom April next year, we’ll start measuring our progress as a country, not just by how our economy is growing, but by how our lives are improving; not just by our standard of living, but by our quality of life.3

Demonstrating that civil servants can also take the lead on incorporating new measures of well-being into policymaking, the Treasuries of Australia and New Zealand have independently developed a Wellbeing Framework and a Living Standards Framework, respectively (Australian Treasury 2006; Treasury 2011). Endowed with the vision of encouraging “higher living standards for New Zealanders” (Treasury 2010, p. 1) since its inception, the New Zealand Treasury finally produced the Living Standards Framework in 2011 to provide guidance on what improving living standards actually amounts to (Treasury 2011). The Living Standards Framework outlines the importance of human, social, and natural capital in addition to traditional economic and physical capital for increasing living standards (Treasury 2011). The Living Standards Framework also sets out the role of subjective measures of well-being as providing a “cross-check of what is important to individuals” (Treasury 2011, p. 1).

Surprisingly, supranational organisations may even end up leading the way on promoting the use of broader measures of well-being in policymaking. In April 2012,

3 Nowhere in his speech does Cameron mention subjective well-being, but the questions, his responses, and the related media coverage of the speech reveal that the most important part of Cameron’s initiative (often referred to as his ‘happiness agenda’) is his request of the United Kingdom’s Office of National Statistics to start measuring subjective well-being. The official transcript of David Cameron’s speech is available from: http://www.number10.gov.uk/news/pm-speech-on-well-being/
the United Nations (UN) General Assembly held a high-level meeting on ‘Happiness and Wellbeing: Defining a New Economic Paradigm’. The agenda for this meeting included the use of measures subjective well-being. Indeed, the World Happiness Report, which was commissioned for the meeting, contained a chapter on ‘The State of World Happiness’ that relies exclusively on research using measures of subjective well-being because “they capture best how people rate the quality of their lives” (Helliwell & Wang 2012, p.11).

Far ahead of the UN, the Organisation for Economic Cooperation and Development (OECD), has identified measures of subjective well-being as essential for a complete understanding of well-being:

For over fifty years, the [OECD] has helped governments design better policies for better lives for their citizens… Ever since the OECD started out in 1961, GDP has been the main factor by which it has measured and understood economic and social progress. But it has failed to capture many of the factors that influence people’s lives, such as security, leisure, income distribution and a clean environment… The OECD Better Life Initiative allows a better understanding of what drives the well-being of people and nations and what needs to be done to achieve greater progress for all… [T]he OECD has identified 11 dimensions as being essential to well-being, [including] overall satisfaction with life… (OECD 2012).4

Lord Richard Layard is the most prominent proponent of the use of measures of subjective well-being, recommending that “quality of life, as people experience it, has got to be a key measure of progress and a central objective for any government”

---

4 The OECD Better Life Index Executive Summary is available here: http://oecdbetterlifefindex.org/wpsystem/wp-content/uploads/2012/02/YourBetterLifeIndex_ExecutiveSummary3.pdf
The Science of Happiness for Policymakers

Weijers

(Layard 2011, no page). Indeed, Layard has even argued that measures of subjective well-being should be the main yardstick for public policy because happiness is the most important goal in life for most of us (Layard 2005, pp. 224–225). Happiness, Layard claims, is “what people want for their children and for their fellow citizens” and thereby concludes that “the greatest happiness of all” deserves to be the ultimate goal of governments and policymakers (Layard 2005, pp. 124–125). Layard understands happiness as meaning “feeling good—enjoying life and wanting the feeling to be maintained” and believes that the emerging field of happiness science has come far enough for us to be able to accurately measure this kind of happiness using subjective survey questions (2005, p. 12). According to Layard, and many others, we should measure happiness precisely because we should be using data from subjective survey measures of happiness to inform policymaking.

Given the reasonable assumptions that sustainably, justly, and equitably increasing the well-being of people should be at least a very important goal for public policy, and given that our subjective judgements about happiness are at least a fairly important part of our well-being, it is clear that policymakers should at least investigate the practicality of measuring subjective happiness. Furthermore, if the citizenry of a democratic state demand that its government includes happiness as one of its overarching goals, then there is overwhelming reason for policymakers to find a way to make measuring happiness feasible.

David’s Cameron’s pronouncement that subjective well-being will be measured in the United Kingdom has been met with mixed comments on blog sites, but the

5 Similar positions are held by Derek Bok (2010), Ed Diener (2011), and Bruno Frey (2008), but none of these authors take the more extreme stance presented by Layard in his (2005) book Happiness: Lessons from a New Science.

6 The use of subjective measures of happiness as the only or ultimate criterion to assess specific policies or progress in general has been criticised by philosophers and economists on many grounds, including most of the traditional philosophical objections to hedonism about wellbeing (e.g. Hausman 2010; Diener & Scollon 2003; Frey & Stutzer 2007). But see also Veenhoven (2010) for a discussion of how the practice of maximising happiness may avoid many of these theoretical problems.
wider public seems to support the notions that happiness should be measured and that the associated results should inform policymaking. In a 2005 BBC opinion poll, 1001 participants were asked whether the government’s main objective should be the "greatest happiness" or the "greatest wealth" and 81% thought that happiness should be the main goal (Easton 2006, no page). This result closely resembles that of a poll taken on The Economist’s website during a debate between Richard Layard and Paul Ormerod, in which the motion “new measures of economic and social progress are needed for the 21st-century economy” received 83% of the support from the online audience (of unspecified size). If these results are representative of popular opinion, then, in democracies at least, we should investigate the practicality of measuring happiness for policymaking. Furthermore, since happiness is ubiquitously understood as being a subjective state, then this means we should investigate the practicality of measuring subjective well-being for policymaking.

1.2 The Basics of Measuring Subjective Well-Being

Whether the use of measures of subjective well-being for policymaking should be pursued depends not only on how important happiness is to people, but also on whether happiness can actually be efficiently and effectively measured. Of all the potential methods of measuring happiness, only questions asking for survey respondents’ judgment about how happy or satisfied they are with their life, are practical on anything but a very small scale. I intend online and smartphone surveys to be included in this. Indeed, as these methods of surveying are becoming cheaper and more accessible, they are being used more and more. Collecting happiness data with behavioural measures, such as expert observations, or any of the neuroimaging techniques, would be prohibitively expensive. Furthermore, it is far from clear that

---

7 The survey results are available here: http://news.bbc.co.uk/nol/shared/bsp/hi/pdfs/29_03_06_happiness_gfkpoll.pdf

8 The audience was well aware that measures of subjective well-being were the new measures in question, since that is all that the debaters and commentators discussed. See the results of the debate here: http://www.economist.com/debate/overview/204
any of these more objective measures are any better at capturing how happy someone is compared to simply asking them. Indeed, the success of neuroimaging measures of happiness is sometimes assessed by the size of their respective correlations with the participants’ responses to subjective well-being survey questions (e.g. Urry et al. 2004).

Subjective measures of well-being can be global or domain-specific. Global measures aim to assess respondents’ judgments of their lives as a whole, while domain-specific measures target limited aspects of respondents’ lives, such as their work lives or their family lives. Although domain-specific measures undoubtedly have their uses, the focus here will be on global measures because they provide a better approximation of the term ‘happiness’ as it is normally understood.

There is a wide range of global subjective well-being questions, but most are subtle variants of general questions about happiness or satisfaction with life. For example, the United States’ General Social Survey asks: “Taken all together, how would you say things are these days? Would you say that you are very happy, pretty happy, or not too happy?” (Kahneman & Krueger 2006, p. 6). The World Values Survey asks: “All things considered, how satisfied are you with your life as a whole these days?”, and uses a response scale ranging from “1 (not at all satisfied)” to “10 (very satisfied)” (e.g. Inglehart et al. 2008). The subtle variations on these questions usually amount to changing the number of points on the response scale or slightly adjusting the wording of the question. For example, the World Values Survey also asks the following question about happiness: “Taking all things together, would you say you are...” with a 4-point response scale: “Very happy... Rather happy... Not very happy... Not at all happy.”

It is widely acknowledged that global subjective well-being questions elicit responses that are biased by contextual factors, the wording of questions, the order and type of preceding questions, and respondents’ current mood (Kahneman &

Experiments have shown, for example, that contextual factors, such as the weather (Schwarz & Clore 1983) and unexpectedly finding a dime (Schwarz 1987), significantly affect how satisfied participants reported being with their whole lives. Experiments on the variability of self-reported satisfaction with life within individuals have demonstrated that people’s reported satisfaction with life as a whole changes dramatically over a period of a few weeks. Kahneman and Krueger (2006, p. 7), for example, found that 218 women who were interviewed twice over two weeks reported life satisfaction scores that correlated only moderately with each other (0.59). Such large differences in how a lot of these women reported judging their life as whole imply that current mood and recent events probably affected their judgments considerably.

Fortunately, large representative samples and careful survey creation can avoid most of these problems. Many of these potential biases can be avoided because they are random biases, which tend to affect different people at different times. By conducting surveys on large representative samples, the impact of random bias on the usability of the results is considerably reduced. In this way, thorough sampling can eliminate the potential bias associated with personal variations in mood, and localised variation in important events (sports teams winning etc.) and the weather. If variability caused by the weather is expected to be a problem, then it can also be reduced by getting participants to acknowledge the weather before completing the survey (Schwarz & Clore 1983). The effects of recent events and participants’ current mood can also be reduced by using a battery of questions about satisfaction with life, instead of just one question (Lucas, Diener, & Suh 1996; Schimmack & Oishi 2005). In order to prevent the significant (but usually small) effect specific questions have on subsequent responses to the more important global questions, researchers usually put global questions first on their surveys (Schimmack & Oishi 2005). Finally, the different results that different wordings of subjective well-being questions produce is a complex issue that is discussed later in this paper.
2. Problems for Using Measures of Subjective Well-Being to Inform Policymaking

Many criticisms have been levelled at the use of happiness science to inform policy and three of the most pertinent will be discussed here: survey measures of happiness are too insensitive, we cannot know what measures of happiness are really measuring, and that the wrong kind of happiness is being measured.

2.1 Are Happiness Scales Insensitive?

Johns and Ormerod (2008) and others have made the claim that time series happiness data is insensitive. While this is a problem for most measures of happiness, it is not an insurmountable one. Johns and Ormerod (2008) argue for the insensitivity of time series happiness data based on a discussion of one particular measure of happiness with a 3-point response scale. This particular selection is somewhat misleading because most recent and contemporary happiness scales have at least four-point response scales, such as the World Values Survey question discussed above. Furthermore, many subjective well-being scales have ten or eleven options.

Put in context, it is easy to see why 3-point response scale might make a measure of happiness insensitive. A 10% increase in average national reported happiness would require 20% (net) of the respondents to respond in the next higher up category (e.g. move from ‘unhappy’ to ‘somewhat happy’ or ‘somewhat happy’ to ‘very happy’). Johns and Ormerod (2008, p. 141) consider such an increase ‘very

---

10 A statistical measure of a dependant phenomenon can be said to be insensitive if its results do not demonstrate statistically significant changes in response to changes of input variables that we have good reason to believe should effect a statistically significant change in the dependant phenomenon.

11 See, for example, the life satisfaction question from the World Values Survey (question V22). See <http://www.worldvaluessurvey.org/> for more information and access to the full data set. See also the well-being questions used in the Gallup World Poll (Gallup Inc. 2008, p. 5). Gallup World Poll questions are available from: <http://media.gallup.com/dataviz/www/WP_Questions_WHITE.pdf>.

12 See Turton (2009, p. 84, note 8) for more details.
difficult” to imagine occurring over “a few years”. This is understandable; a 10% increase in reported happiness over a few years will be unusual no matter what measure of happiness is used. Consider that, to achieve a 10% increase in reported happiness on a 10-point scale, 90% (net) of respondents would need to report a higher category of happiness or 22.5% (net) would have to report their happiness as four categories higher (Turton 2009, p. 85).

A further problem exists, however, for 3- or 4-point response scales: They are likely to be insensitive on the micro (individual person) level because a person would have to undergo a considerable change in their judgement of their life to move at all on the response scale. So, the worry with 3- or 4-point response scales is that the big changes in judgments of happiness required to change the results on the micro level, will make observable changes at the national level more unlikely than when a 10- or 11-point response scale is used (Cummins & Gullone 2002). While this is undoubtedly true, a statistically significant 10% increase in average reported happiness (on 4-point scales) has occurred in some countries over just a few years, including Lithuania (1997–1999), Mexico (1996–2000), and Slovenia (1992–1995), and in many more countries, including Johns and Ormerod’s home country of Great Britain have reported a 10% or greater increase in reported happiness over slightly longer periods of time (Ingleheart et al. 2008; Turton 2009, p. 85).

Clearly then, while time-series happiness measures might be relatively insensitive, even happiness measures with 4-point response scales are sensitive enough to capture trends. Therefore, worries about the insensitivity of happiness measures are not completely unfounded, but they do not provide a good reason to avoid using time series happiness data to guide policy. Indeed, many time series happiness studies can be useful for policymaking in several ways (Frey 2008, Chap. 13). For example, careful comparison of survey data from similar nations, or other groups, where policy change has occurred in some groups but not others, can help to evaluate the effects of policy change on reported happiness (Turton 2009). This method works best when detailed datasets are available because they help to isolate the effects of the policy change from those of other factors that are known affect reported happiness.
2.2 Do Measures of Happiness Really Measure Happiness?

As discussed, there are currently a number of different methods that are claimed to be measures of well-being. These measures include brain scans, daily reports about how participants have been feeling, the opinions of participants’ friends or colleagues, the opinion of an expert, the amount participants smile and, most commonly, survey questions (Weijers 2010). Importantly, there is a range of survey questions about happiness and well-being, many of which are worded very differently. Some survey questions ask about positive and negative feelings, satisfaction with life, whether you would change anything about your life, and, of course, happiness. These diverse survey measures, collectively referred to as subjective well-being or happiness measures, are what policymakers are considering using to guide their policy decisions. But should we trust that these surveys are actually measuring happiness or well-being and not something else entirely? And should we trust that one person’s happiness is the same as another’s?

Kroll (2010), Layard (2003; 2005) and Bok (2010) all argue that we should have faith in answers to survey questions about well-being because they are significantly correlated with many other measures of well-being, both within and between large groups of people. Importantly, these researchers view the correlations found in (the unarguably objective) neuroimaging studies as adding considerable support to the idea that subjective measures of well-being are assessing something that is real and that we all experience. Layard (2005, p. 17) states:

Sceptics may still question whether happiness is really an objective feeling that can be properly compared between people. To reassure doubters, we can turn to modern brain physiology with its sensational new insights into what is happening when a person feels happy or unhappy.

Indeed, the diverse measures of well-being mentioned above are nearly always statistically significantly correlated (Weijers 2010). However, that truth does not necessarily entail that any of the different measures are actually evaluating well-being or happiness. The important point that Kroll, Layard, and Bok overlook is that,
although the majority of the correlations between neuroimaging and subjective well-being survey data are statistically significant, they are not significant in size (Weijers 2010; Weijers & Jarden 2011). Furthermore, the neuroimaging studies that Layard (2005, pp. 17–19) and others cite usually compare neuroimaging results with various cues that are expected to create various feelings, including ‘approach’ or ‘withdrawal’, in their participants, not with measures of subjective well-being.

In the only comprehensive study of correlations between neuroimaging and measurement of subjective well-being, several prominent subjective well-being measures were compared with electroencephalogram data from 84 right-handed adults aged 57–60. Correlating highest with the neuroimaging results were the results for a measure of psychological flourishing—Carol Ryff’s (1989) Scales of Psychological Well-Being. The correlation was highly significant \((p < 0.01)\) and moderate in size \((0.33)\) (Urry et al. 2004, p. 370). Following close behind was Ed Diener and colleagues’ (1985) Satisfaction With Life Scale, which correlated with the neuroimaging results by \(0.30\) and was also highly significant \((p < 0.01)\) (Urry et al. 2004, p. 370). A smaller \((0.21)\) and slightly less significant \((p < 0.05)\) correlation was also found with the positive affect component of Watson, Clark, and Tellegen’s (1988) Positive and Negative Affect Schedule (Urry et al. 2004, p. 370). It should be noted that, unlike the questions that directly ask about happiness, none of the measures of subjective well-being used in Urry and colleagues’ (2004) study are equivalent to the folk notion of happiness (although the measure of positive affect would come the closest).

The significant, but relatively small, correlations between the neuroimaging results and these measures of subjective well-being tell us two things (Weijers 2010). First, the neuroimaging results and these measures of subjective well-being are very likely to be tracking phenomena that are related in some positive way. This is shown by the fact that the correlation is positive and highly statistically significant. Second, we can be confident that the phenomena being tracked are distinct. The high statistical significance of the results should make us confident that the various measures are not measuring exactly the same thing. If the size of the correlations
were much higher, at least above 0.60, and the statistical significance remained very high, then we would expect to observe the phenomena measured by the different tests to covary more closely and, thereby, give the impression of being the same thing. Positive correlations of 0.33 (roughly) mean that we should expect an increase in the results of the neuroimaging measure to be usually accompanied by a relatively smaller increase of the subjective well-being measure. This is the kind of relationship we expect from distinct but positively related variables, not from two different measures of the same phenomenon.

Perhaps most telling of all on the question of what support neuroimaging provides for the objectivity of happiness is that the cognitive scientists who carry out neuroimaging studies rarely claim to be testing happiness or well-being. Much more commonly they claim to be investigating the neural correlates of pleasure and pain or approach and withdrawal behaviour, as is the case with the neuroimaging study that Layard discusses the most (Davidson et al. 1990; c.f. Layard 2005). But even if experimental neuroimaging studies were carried out until a measure of neurological activity correlated very highly and statistically significantly with a subjective measure of happiness, it would be presumptuous to declare it the discovery of an objective measure of happiness. Rather, such a neurological measure should be understood as an objective measure of the propensity to report subjective happiness. Understood this way, it’s much less obvious how objective neuroimaging results are supposed to give us confidence that a measure of subjective well-being is accurately performing the task that we want it to.

Nevertheless, the confluence of the various correlations between measures of subjective well-being, aspects of bio-physical health, neuroimaging data, observers’ reports, and behavioural analyses points toward there being something tangible to

---

13 0.60 is the level above which behavioural scientists usually deem results to be ‘highly related’ (Cohen 1988).

14 A similar, but less detailed, presentation of this issue appears in Weijers & Jarden (2011, pp. 56–57) and Feldman 2010, chap. 13).
measure (Layard 2010). Or, at least, a few closely related tangible things to measure.\footnote{Indeed, this is surely how these results are best understood. The fact that different measures of well-being actually measure different \textit{aspects} of well-being is not necessarily a drawback, as discussed below.} Frey and Stutzer (2002) also provide a (now dated but nonetheless excellent) summary of how measures of subjective well-being are reliable \textit{enough} to provide useful economic and policy insights despite all of the potential problems discussed so far.\footnote{For a more up-to-date, but less comprehensive review, see Helliwell and Wang (2012, pp. 17–19).} Perhaps most encouraging in this regard is the study of Oswald and Wu (2010, p. 579), which reported a highly significant ($p$ approximately $= 0.0001$) and relatively large (0.6) correlation between objective quality of life factors and subjective satisfaction with life in the United States of America. This careful study of over 1.3 million data points “suggests that subjective well-being data contain genuine information about the quality of people’s lives”, which is exactly what policymakers should be interested in (Oswald & Wu 2010, p. 579). A further problem remains, however. None of the many existing measures of subjective well-being measure the same thing, so it is far from obvious which one or ones, if any, actually measure well-being.

### 3.2 How Do We Know if We Are Measuring the Right Kind of Happiness?

That the various questions in well-being surveys are not tracking the same phenomena is well known by happiness researchers. Indeed, many theoretically inclined social scientists and philosophers recognize that the different kinds of questions used in subjective well-being surveys often endorse one particular theory of well-being. Do these differences matter if all of the measures are assessing something that is obviously good? In many policy situations, yes they do.

What if findings based on different measures of well-being imply different policies (Weijers 2010)? This is not just a theoretical problem, as many seemingly contradictory results from ‘happiness science’ show. For example, Kroll (2010) states...
that it is received wisdom amongst happiness researchers that increases in annual income (over about 10,000 Euros) have no effect on the average reported happiness in country. However, other happiness researchers have found a statistically significant relationship between income and reported happiness in rich as well as poor countries (e.g. Deaton 2010). But this is no contradiction (and so no reason to think that happiness research is unreliable) because Deaton and the other happiness researchers who do find a statistically significant relationship between income and reported happiness use different measures of happiness. In Deaton’s (2010) case, the finding that increasing income is correlated with reported happiness well above 10,000 Euros is based on data from Gallup World Polls. The Gallup happiness question is worded in a way that seems to encourage respondent to make comparisons with all other people in the world, not just their immediate reference group (Weijers 2010). Many studies have shown that our reported satisfaction with life is significantly affected by whatever reference group is most salient to us at the time of the survey (Graham & Pettinato 2002; Kahneman & Krueger 2006; e.g. Merton 1957). Therefore, it should come as no surprise that Gallup World Polls’ wording of their subjective well-being question makes a difference in this way—people in wealthy countries feel better about their lives when they think about how poor people in other countries are.

Indeed, other researchers have good evidence that the more a measure of subjective well-being asks about the respondents’ emotional lives, and the less it encourages them to engage in cognitive deliberation about how satisfied they are with their life compared to what it might have been, then the smaller and less significant the relationship between increases in income and increases in subjective well-being becomes (Diener et al. 2010).

So, how can we deal with the problem of apparently contradictory findings from the science of happiness? First, happiness researchers should avoid generalising findings from different measures of well-being, unless those questions have been shown to track the same phenomena in the same circumstances in other studies. Second, policymakers should always refer to the original research papers and even the survey questions themselves if the researchers have not made it obvious (Weijers
Having found that the contradictory happiness science findings are based on the use of different measures of well-being, what should a policymaker do? Discovery of such a conflict illuminates the fundamental problem policymakers intent on using happiness science face: Which question about subjective well-being is the most appropriate basis for policymaking (Weijers 2010)?

Policymakers should not expect to have to answer this question by themselves. Philosophers have debated the merits of various theories of well-being for at least two thousand years and social scientists have been arguing over which measures of well-being are the most reliable and the most representative of the best philosophical theories of well-being. Unfortunately, and despite considerable efforts, there is still no agreement between academics on which question about subjective well-being is the most appropriate basis for policymaking. But, thanks to these academics, there are many candidates, whose advantages and disadvantages have been discussed at length. This academic knowledge needs to be discussed widely to engender public debate on what makes people’s lives go well for them and the proper aims of government. A populace that is informed about what makes citizens’ lives go well for them will be able to exercise their democratic rights to lobby and perhaps vote for their preferred conception of wellbeing (Weijers 2010). When this occurs, policymakers can work with social scientists to ensure that appropriate measures of well-being are used to guide public policy. This will allow happiness researchers to be confident that they are producing findings that are relevant for policymaking and policymakers to fully benefit from happiness science.

4. What Role Should Happiness Play in Policymaking?

In any democratic society the (hopefully informed) citizens should decide what conceptions of happiness or well-being are important and the extent of the role any

---

17 Dan Haybron and Valerie Tiberius provide some excellent examples of how philosophers can engage with social scientists and policymakers to inform them about the philosophical theories and debates about well-being (Haybron 2008; Haybron & Tiberius 2012; Tiberius 2004; Tiberius & Hall 2010).
such conceptions should play in policymaking. But in order to educate citizens and encourage effective evidence-based policymaking, academics and top-level civil servants need to better clarify the various conceptions of happiness and well-being and whether we can accurately and efficiently measure them. After these issues have been clarified, how circumstances and policies affect happiness or well-being (as defined in each case) should also be investigated to help better understand where each concept of well-being fits in the economic landscape, and also to promote public debate on the relevant merits of certain kinds of happiness and other goods. Philosophers, psychologists, economists, statisticians, political scientists, and policymakers should work together on this in order to pool their collective expertise and progress most effectively. But what can governments and policymakers do to speed up this process?

Nobel Laureate Daniel Kahneman has voiced his worry that governments will import only one measure subjective well-being, such as the satisfaction with life question, in their censuses (Jarden 2011). He fears that this kind of measure could be too insensitive to be able to significantly correlate trends in reported happiness with changes in policy. Of most concern to Kahneman is if policymakers will see the insensitivity of such measures as a reason to abandon happiness science entirely. So, how can governments and national statistical offices make sure that they put the right question in the census from the start?

In fact, this question is considered a red herring by most happiness researchers. As Martin Seligman discusses in *Flourish* (2011), one question is simply not enough. Indeed, there is considerable evidence in favour of a ‘dashboard’ approach—the use of several headline indicators of well-being (that may be based on individual questions, but would be better based on an index of several related questions). The ‘dashboard’ approach is popular amongst happiness researchers because philosophical and social scientific theory generally supports the notion that there are several distinct and roughly equally important dimensions of well-being (or at least people disagree about which of the different dimensions are important) (Weijers 2011). Seligman argues that positive emotion, engagement, meaning, positive
relationships, and accomplishment are worthy of inclusion in a dashboard of subjective well-being indicators (Forgeard et al. 2011; Seligman 2011).

The United Kingdom’s Office for National Statistics (ONS) appears to have followed the dashboard approach to some extent, asking four subjective well-being questions in its ONS Opinion Survey, each of which represents a different group of philosophical theories about well-being (Beaumont 2011). The ONS has not yet confirmed which questions will capture data on subjective well-being in the long-run, with consultation still ongoing. It seems very likely, though, that several subjective well-being questions will be chosen. This cautious approach is surely a good one. Indeed, more research is required in order to better understand the most important aspects of well-being and the best way to measure each of them given the usual constraints on governmental data collection (Weijers 2011). So, the question should be: how can governments and national statistical offices make sure that they put the right questions in the census from the start?

A ten-year international collaborative effort could go a long way to answering this important question. Several major longitudinal panel surveys are already conducted around the world, and some of these surveys ask subjective well-being questions. The gold-standard of these surveys for our purposes are those in which each respondent, and everyone else in their household, completes the survey periodically for many years (Weijers 2011). These existing panel surveys would have to be standardized to some extent, including ensuring that all of them asked questions about the events that have recently impacted their lives and a battery of subjective well-being questions for each potentially important aspect of well-being18 (in addition to the traditional economic and demographic questions.19

---

18 Comparing the various measures of subjective well-being to the existing philosophical theories of well-being would be a useful way to assess whether the existing measures cover the scope of all possibly important conceptions of well-being. Such an assessment would have to be carried out carefully because of the profound differences between philosophical accounts of well-being that are nonetheless grouped together. For example, there is a variety of types of hedonism about well-being. While a simple “How happy have you been these days?” question might be thought to cover all hedonistic theories, other measures better
If this kind of survey could be conducted in several countries around the world, happiness researchers could achieve a lot with the results (Weijers 2011). Happiness researchers could assess what kinds of objective economic and demographic factors affect the various aspects of well-being and if these effects are cross-culturally robust. By surveying the same participants, and those they live with, over time, the surveys will help happiness researchers better understand the complex issues of adaptation to life events and the relativity of reported happiness to the various attributes of the people around them (Clark forthcoming).

Furthermore, if the surveys were large enough, sub samples of the surveyed populations could be given slightly altered questions. This would enable happiness researchers to test the accuracy and sensitivity of different questions. At the conclusion of the multinational trial, governments and policymakers could decide which measures of well-being are best suited to their respective citizenries and policy goals (Weijers 2011). Indeed, the OECD’s current work on creating guidelines for measuring subjective well-being, combined with the political impetus created at the UN meeting on Happiness and Wellbeing: Defining a New Economic Paradigm held in April this year, might result in such a widespread collaborative effort.

If well-being dashboards were incorporated into censuses, then political parties of the not-too-distant future should be able to differentiate themselves by claiming that their policies favour certain aspects of well-being over others (Weijers 2011). To better facilitate public debate on the merits of these various measures of well-being, philosophers, psychologists and economists should hold interdisciplinary public capture the ideas behind specific hedonistic theories of well-being. For example, Kahneman and Krueger’s (2006) U-index (which asks about the proportion of time respondents spend in a negative emotional state) would be a fairly good measure of Epicurus’ pain-minimising hedonism, while Watson and colleagues’ (1988) Positive and Negative Affect Scale would better represent Bentham’s (1789) net-pleasant-feelings-based hedonism.

Ed Diener’s (2006) ‘Guidelines for National indicators of Subjective Well-Being and Ill-being’ is a good starting point for more specific guidance on the creation of such a collection of measures of subjective well-being.
lectures and forums on different conceptions of well-being. Assuming that the public gain sufficient knowledge about the different aspects and conceptions of well-being being measured, the constant collecting of data on these well-being indicators could be a great way to measure a government’s effectiveness (comparing the net benefit to well-being indicators with the net cost to capital stocks, such as natural resources, infrastructure, etc.) (Weijers 2012). In this way, measures of subjective well-being could be used to allow policymakers to consider the potential impacts of a policy on the happiness as well as the wealth of citizens. It will take a considerable amount of interdisciplinary work to reach this point but, once there, policymakers will have an incredibly useful set of tools at their disposal and citizens might just become happier because of it.

References


---

20 Such as the June 2012 Wellbeing and Public Policy Conference held in Wellington, New Zealand (http://www.wellbeingandpublicpolicy.org) (Morrison & Weijers 2012).


