USING HAPPINESS TO VALUE HEALTH

PAUL DOLAN

OHE Research
Office of Health Economics

USING HAPPINESS TO VALUE HEALTH

PAUL DOLAN

NOVEMBER 2011

Office of Health Economics,
Southside, 7th Floor, 105 Victoria Street, London SW1E 6QT
www.ohe.org

ABOUT THE AUTHOR

Paul Dolan is Professor of Behavioural Science in the Department of Social Policy at the London School of Economics and Political Science. He also has held positions at Imperial College Business School and the Universities of Sheffield, Newcastle, and York and has been a visiting research scholar at Princeton University. He has a DPhil in Economics from the University of York. His professional roles currently include Chief Academic Advisor on Economic Appraisal for the UK Government Economic Service, a member of the Measuring National Wellbeing Advisory Forum for the Office for National Statistics in the UK, and a member of the National Academy of Science panel on subjective wellbeing in the US.

Prof Dolan's research career encompasses four related research activities: the valuation of health to generate quality-adjusted life years (QALYs); developing ways to weight QALYs that account for concerns for fairness; developing measures of subjective wellbeing for policy purposes; and advancing the use of behavioural economics to address policy challenges. It is these activities, alongside the influence of working with Danny Kahneman at Princeton, that have provided the motivation for this monograph.

ACKNOWLEDGEMENTS

I am grateful to Pfizer, particularly Julie Anne Bridge and Warren Cowell, for funding previous work, some of which is reproduced in this report. I owe a huge thank you to Henry Lee for helping me to shape this report and to Robert Metcalfe and Tessa Peasgood, who worked with me on the previous report and who are responsible for the data analysis reported here. Two referees provided incredibly helpful comments that have served to improve the report. I would also like to thank Richard Suzman for funding the US study and Arie Kapteyn and those at RAND for gathering the data for that study. Data from the British Household Panel Survey (BHPS) were supplied by the UK Data Archive. Neither the original collectors of the data nor the archive bear any responsibility for the analysis or interpretations presented here. The views expressed here are my own and I take full responsibility for any errors.

ISBN: 978-1-899040-14-8

CONTENTS

EXECUTIVE SUMMARY

Context

- Existing preference-based methods for valuing states of health and illness present problems
- Interest in happiness research is increasing and the UK is at the forefront of this
- Happiness is defined in very broad terms, encompassing not only experiences of mood but also evaluations of life satisfaction
- Health, particularly mental health, is a major determinant of all forms of happiness

Objectives

- Consider the current ways of valuing health and some problems with existing preference-based methods, as motivation for the happiness-based approach as an alternative
- Consider happiness research and its applicability to valuing health, for the purposes of showing what matters to people and, ultimately, for use in economic evaluation
- Analyse the British Household Panel Survey (BHPS) and new data from the US to generate happiness-based tariffs for the SF-6D and EQ-5D; compare these tariffs with those estimated using standard gamble and time trade-off preferences elicited from the UK general population
- Discuss the implications for the field of happiness and health policy in light of these new analyses and suggest next steps for policy makers

Results

- The dimensions of health privileged by the EQ-5D and SF-6D may not be those that most affect people's lives
- Of the domains of health within the SF-6D and EQ-5D, mental health and vitality appear to be most strongly associated with happiness, whilst physical functioning and pain are not so strongly associated with happiness
- Preference-based tariffs show that physical functioning and pain matter as much to people, and sometimes more, as mental health when they are asked to risk death or trade off life years

Implications

- Happiness data allow us to say what is important in people's lives when they are not thinking about how important those things are, and it obviates the need to be prescriptive about what dimensions of health matter to people
- The dimensions of health that matter most in happiness regressions are not the same as those that matter most when people are asked about their preferences
- Mental health matters more in happiness reports and so one implication of using happiness to value health is that greater priority would be given to mental health as compared to physical functioning and pain
- Further research is required to test the robustness of these findings and to better understand the source of the difference between happiness-based and preference-based valuations for health

1. INTRODUCTION

Interest has surged in "wellbeing" in recent years, both in popular culture and in the science of "subjective wellbeing" or "happiness research". Perhaps most significantly, policy makers have embraced the idea that measuring our happiness can provide meaningful information that can, in various ways, be fed into decisions about how to monitor progress, and to inform and evaluate policy (Dolan, et al., 2011). The recent and influential report commissioned by French President Nicolas Sarkozy recommends that "Measures of both objective and subjective wellbeing provide key information about people's quality of life. Statistical offices should incorporate questions to capture people's life evaluations, hedonic experiences and priorities in their own survey . . . [R]esearch in recent years has enabled us to improve our metrics, and it is time to incorporate in our measurement systems some of these advances" (Stiglitz, et al., 2009).

The academic and philosophical tradition of interest in happiness is lengthy. Jeremy Bentham described utility in terms of pleasure and pain (Bentham, 1789) and Francis Edgeworth described a hedonimeter to display these experiences over time (Edgeworth, 1881). As a result of the behaviourist revolution in the social sciences, and the belief that happiness could not be measured and interpersonally compared, economists turned to a new interpretation of utility in terms of "wantability" (Fisher, 1918). The trend to measure utility through observed behaviour gathered pace and hedonic accounts of wellbeing became less popular. A preference-based account of welfare lies at the heart of economists' attempts to value health, but it is not without problems. The roots of what we currently think of as happiness research are grounded in the work of psychologists over the last 40 years or so (Diener, 1999) and, more recently, by economists who have questioned the reliance of observed behaviour as a guide to wellbeing (Dolan and Kahneman, 2008).

This report provides an accessible guide into the latest developments in happiness research as they apply to the valuation of health. It presents new valuation data that show how the dimensions of health that matter most in happiness regressions are not the same as those that matter most when people are asked about their preferences. In particular, mental health matters more in happiness reports. Therefore, one implication of using happiness to value health would be that greater priority is given to mental health as compared to physical functioning and pain. This report considers the degree to which data of the kind reported here can overcome some of the well-known problems with existing preference-based ways of valuing health for the purposes of informing resource allocation decisions.

It is worth emphasising that I focus on issues pertaining to how best to value health benefits for the purposes of informing resource allocation decisions. As such, I do not consider the decisions of private individuals, where preferences have much greater resonance and legitimacy. I certainly make no prescriptions about how individuals should spend their own cash, although such decisions also usefully could be informed by the consequences for happiness. I also do not consider wider issues concerning the objectives of the health system. I fully appreciate that benefits will be considered alongside other, sometimes competing, concerns relating to inequalities in health, access to and so on.

1. INTRODUCTION

This report has four main objectives:

- Consider the current ways of valuing health and some problems with existing preference-based methods (which rely on the degree to which people are willing to trade future life years or risk of death for future improvements in health) as motivation for the happiness-based approach (which relies on direct assessment of life satisfaction or daily mood) as an alternative
- Consider happiness research and its applicability to valuing health, for the purposes of showing what matters to people and, ultimately, for use in economic evaluation
- Analyse the British Household Panel Survey (BHPS) and new data from the US to generate happiness-based tariffs for the SF-6D and EQ-5D (regressing these descriptions of health on happiness ratings), and compare these tariffs with those estimated using standard gamble and time trade-off preferences (regressing the same descriptions of health on preference assessments)
- Discuss the implications for the field of happiness and health policy in light of these new analyses and suggest next steps for policy makers

The following four chapters deal with each of these issues in turn.

2. CURRENT WAYS OF VALUING HEALTH AND SOME PROBLEMS

Key points:

- Economists value health according to the willingness of members of the general public to risk of death or trade off life years for different health states defined in ways prescribed by classification systems such as the EQ-5D
- Problems exist with narrow descriptions of health and especially with hypothetical valuations, which often bear little relation to the real impact the health states have
- We need to judge preferences according to a substantive account of rationality based on the consequences of different choices

Markets value goods and services (and treatments) according to our revealed preferences, expressed through our willingness to pay (WTP). We usually do not pay for other "non-market" goods, however, at least not in the same way we would buy a car or a pair of shoes. It is still helpful for policy purposes to value these non-market goods in monetary terms so that we can compare them directly to costs. Monetary values for benefits are derived through WTP estimates, either by inferring values from preferences revealed through market behaviour or stated directly in contingent valuation studies. The Department of Treasury in the UK provides guidance on these approaches in the Green Book (HM Treasury, 2003). A notable exception is health benefits, which often are valued in health, not monetary, units. The demand for non-monetary values came about partly as a "push" against WTP and partly as a "pull" from evaluations in health care. The stated WTP values for non-market goods have been found to be sensitive to factors to which they should be insensitive, such as the size of the risk reduction. The appropriateness of using WTP to value health also raises ethical concerns, for example, because value will be related to ability to pay (and the income distribution may be seen as inequitable) and because of the signal that using money to value health may send (implying health is just like any other commodity). These considerations have given rise to a desire to find alternative metrics.

Quality-adjusted life years (QALYs) have been designed as this metric, combining the two main outcomes of health care: quality and length of life. QALYs attach quality of life weights to different states of health and illness and then multiply those weights by how long the states last. The "Q" in the QALY is calibrated on a cardinal scale between 0 (for dead) and 1 (for full health) so that it is possible to say how much better or worse one state is compared to another. One QALY represents one year of life in full health, or two years in 0.5 health, and so on. QALYs cannot be used to determine the resources that should be devoted to health care compared to other areas of public policy, but they can be used to allocate health care resources. An efficient health care system will be one that invests in interventions that generate the most QALYs at least cost. Allocations based on QALYs would effectively work down a "league table" that ranks all interventions in terms of their incremental cost-per-QALY ratio until health care resources have been exhausted.

The National Institute for Health and Clinical Excellence (NICE) has led the way in the appraisal of health interventions and technologies in the UK, setting out its preferred QALY-based methods in its "reference case". Other countries are following NICE's lead (Chalkidou, et al., 2009). It is my contention that current practice and received wisdom in health economic evaluation is too narrow in the way it thinks about benefits (in relation to both how benefits are defined and whose benefits are included). In line with the spirit of the Green Book (but not its specific recommendations in relation to WTP), our approach "emphasises the need to take account of the wider social costs and benefits of proposals, and the need to ensure the proper use

of public resources" (UK Treasury, 2003). Note that I focus on issues pertaining to how best to value health benefits for the purposes of informing resource allocation decisions and not on wider issues concerning the objectives of the health care system. I fully appreciate that benefits will be considered alongside other, sometimes competing, concerns relating to inequalities in health, access to health care and so on.

The QALY approach requires that we both describe quality of life and value it. In fact, we must answer three core questions: what is to be valued, how it is to be valued and who is to value it (Dolan, 2000). In terms of description, generic descriptive systems have been designed for calculating QALYs. NICE currently recommends the EQ-5D (EuroQol, 1990), which defines health in terms of five dimensions (mobility, self-care, usual activities, pain and mood) with three levels of severity for each dimension; thus generating 243 (35) possible health states. The SF-6D is another contender (Brazier, et al., 2002), collapsing the SF-36 into six dimensions – physical, role and social functioning, pain, mental health and "vitality" (encompassing notions of energy and zest for life) – with four to six levels of severity (resulting in 18,000 health states).

Values for these health states – the value of the "Q" in the QALY – are derived from people's preferences for different health states. These preferences are elicited through expressed willingness to take a risk of death, the standard gamble (SG) method, or to give up life years for improved health, the time trade-off (TTO) method. The SG requires respondents to consider the combination of the risk of full health (p) and the risk of death (1-p) that is equivalent to the certainty of a poor health state. The TTO requires respondents to consider how many years of life in full health, x, are equivalent to a longer time, t, in a poor health state. If full health is assigned a value of 1, then the value of the poor health state is taken to be p for the SG and x/t for the TTO. NICE currently favours the TTO.

For the EQ-5D, 3395 members of the general public were initially asked to rank 13 EQ-5D states[1], plus immediate death and unconsciousness. They then were asked to rate those states on a 0-100 scale and then to value them using the TTO method. The time in the poor health state was set at ten years; the respondent was asked to state the number of years in full health equivalent to this. If the respondent said immediately death was preferable, then the protocol for states worse than dead was to set immediate death equivalent to a scenario involving some fraction of ten years in the state rated as worse than dead, followed by the remainder of those ten years in full health (Dolan, et al., 1996). From 2997 respondents with complete data, a tariff of values for all 243 states has been estimated (Dolan, 1997).

In the case of the SF-6D, the values attached to the 18,000 states generated by the classification system were derived from SG valuations for a sample of 249 of these health states. Face-to-face interviews were conducted with a representative sample of 611 members of the UK population (Brazier, et al., 2002). Respondents initially ranked five SF-6D health states, plus the best and worst states and immediate death. The SG questions then asked respondents to choose between each of five certain SF-6D states (imagining remaining in those states for the rest of their lives), versus a gamble between the best and worst health states. Respondents next were asked to value the worst state in relation to immediate death. These valuations were revised in 2008 partly to deal with missing item level data, and partly to address some of the inconsistencies in the original tariff values (Brazier, et al., 2008).

[1] This always included descriptions of full health (health state 11111) and also the worst possible health (health state 33333) as described by the EQ-5D.

Critically, in relation to the question of "who", NICE asks for values to be elicited from the general public in the form of hypothetical preferences, as opposed to individuals who have experienced these specific health states. Using the preferences of the general public is broadly consistent with an insurance principle, whereby the ex-ante preferences of those who might be affected by a condition in the future are given weight in allocating resources (whether they should be given all the weight is another matter entirely and to which we shall return later). These recommendations are also in line with those proposed by the Washington Panel (Gold, et al., 1996).

These recommendations about what is valued, how and by whom, raise some serious concerns. In recommending the EQ-5D, NICE and other agencies are saying that the dimensions of health within it are the only important ones; no good normative or empirical basis for this claim is evident, particularly when other dimensions, such as vitality (shown consistently to be an important domain of health yet not included within the EQ-5D [Grieve, et al., 2009]), affect people every bit as much as the '5D' in the EQ-5D. Since the EQ-5D was designed for use among patient populations, it also is not clear that the EQ-5D picks up the impact that conditions have on the families of patients. The impact on others affected by the condition is increasingly recognised as an important consideration in health technology assessments (e.g., in recent discussions about the cost-effectiveness of treatments for Alzheimer's) and we need to do more to accurately capture those effects.

Responses to methods, such as the SG and TTO, are subject to various biases that mean they rarely reflect real experiences. In many ways, our health state preferences more accurately reflect our affective reactions to – or fears about – particular health states rather than considered assessments of what life would actually be like in those conditions. Such considered assessments should account for any adaptation, which I define here as adjustment to new or changed circumstances, where the adjustment is within "oneself" rather than in relation to the use of corrective devices such as spectacles or prostheses. Whilst fear is certainly an important thing to consider in health care, it is not what the SG and TTO are designed to tap into; rather, they are intended to reflect a cognitive assessment of the impact that a particular health state will have on utility.

In any event, the fundamental problem (from whomever preferences are elicited) is that the things that we focus on in a preference question are often not the things that we focus our attention on and that matter to us in the actual experiences of our lives (Schkade and Kahneman, 1998; Gilbert and Wilson, 2000). The general public tends to overestimate the severity of the loss from many (but not all) health conditions (De Wit, et al., 2002), partly because it exaggerates the extent to which patients attend to their health state (Dolan and Kahneman, 2008). Imagine being asked to value walking with a cane. It is almost impossible to avoid imagining that as you walk you will be thinking about the cane much of the time when, in fact, the cane will rarely be the focus of your attention, especially as time passes. Such focusing effects are an issue for any preference elicitation question for any population, including "patients", since what we focus on in the question may not be focused on to the same extent in the experience of our lives. A person who walks with a cane and is asked to imagine having their walking restrictions alleviated will inevitably imagine actively enjoying the freedom of normal walking, which they would then quickly take for granted.

A reliance on the EQ-5D, the TTO and the general public means that NICE and other agencies valuing QALYs using preference-based methods may make the wrong decisions about which treatments to recommend: Wrong in the sense that more benefit could be gained by making different decisions. More benefit to me means better experiences, greater happiness. Utility as preference satisfaction and happiness maximisation have the same extension if people want what they will eventually enjoy. If health state A is preferred to health state B, then A should produce the greatest flow of happiness.

Now, this may not be what others have in mind when they think about greater benefit: They may simply care about preferences being satisfied, however perverse or pathological those preferences. Most economists, however, do impose coherence constraints and preference elicitation studies in health and, elsewhere, increasingly have in-built checks for logical consistency, transitivity, and test-retest reliability. These are to assess whether an individual's preferences are self-contradictory according to some, typically axiomatic, model. By expressing coherence as a set of axioms, it is possible formally to define a "rational" economic agent and to derive testable predictions about the class of future (rational) choices that would be consistent with past (rational) choices.

On the face of it, economists have been reluctant to require that preferences satisfy a more substantive account of rationality that would establish some way of assessing outcomes as they occur, not only as they are conceived at the time of the decision. Almost all preference satisfaction accounts, however, require that our choices are based on "informed" desires (Harsanyi, 1985). To me, "informed desires" make accurate predictions about the future utility associated with different choices. It is unrealistic to assume that preferences will be sufficiently informed to maximise future utility without having first established whether or not this assumption holds, and so we should at least enquire into where and how what we want most differs from what we enjoy best. In any case, "preference in relation to experience" is adopted implicitly in many economic analyses, explaining why the terms "utility", "preferences", and "happiness" often are used interchangeably in economics textbooks.

3. HAPPINESS RESEARCH AND ITS APPLICABILITY TO VALUING HEALTH

Key points:

- Interest in happiness research is increasing and the UK is at the forefront of this
- Happiness is defined in very broad terms, encompassing not only experiences of mood, but also evaluations of life satisfaction
- Happiness data allow us to say what is important in people's lives when they are not thinking about how important those things are, and it obviates the need to be prescriptive about what dimensions of health matter to people
- Health, particularly mental health, is a major determinant of all forms of happiness

Many economists recognise some of the problems with preference-based approaches and are looking for ways of refining conventional QALY methods. My approach involves looking for more suitable ways of capturing the real experiences associated with the use of new therapies, capturing the whole experience of a health care intervention, and accounting for the impact on the family as well as on the patient. If we could find a suitable metric that would allow us to make these various comparisons, we would be much nearer to being able to join up quality, patient satisfaction and cost effectiveness (Darzi, 2008). This has led me to focus on happiness (Dolan and White, 2007). It does not necessarily require us to abandon the QALY approach of weighting each health state by its duration. Indeed, some assessments (e.g., based on daily reports of happiness) relate to short and specific periods of time and so lend themselves directly to duration weighting. A focus on happiness, however, does allow us more flexibility in how health is described and it allows us to think about valuation in a more complete way than focusing only on SG- and TTO-type preferences. It also has considerable potential to be incorporated into the Value Based Pricing (Department of Health, 2010) methodology for new pharmaceuticals within the UK NHS.

The many attempts to classify the different ways in which happiness can be measured fall broadly into "evaluation" and "experience" (Kahneman and Riis, 2005). Happiness is measured as an evaluation when people are asked to provide global assessments of satisfaction with their lives overall. Happiness measures are elicited from the public, or participants in clinical trials, in the form of paper or electronic questionnaires or telephone interviews. Economists have been using "life satisfaction" for some time (see Frey and Stutzer, 2002). Experience is closely associated with the Benthamite view of wellbeing, where pleasure and pain are the only things that are good or bad for anyone. This may be colloquially thought of as the amount of "affect", a term psychologists use for emotion, felt in any moment (e.g., happy, anxious, etc.).

Evaluations and experience-based measures may sometimes produce similar results (Blanchflower, 2009), but often do not. For life satisfaction, it appears that unemployment is very bad; marriage is pretty good, at least to start with; children have no effect; retirement is pretty good, at least to start with -- but there is considerable heterogeneity in these assessments (Calvo, et al., 2007). In contrast, associations between affect and these events generally are weak (Knabe et al, 2010). However, thorough comparison of the measures of happiness has not been done because no large scale longitudinal or repeated cross-sectional survey containing all measures has been done. This is about to change in the UK as the Office for National Statistics, with the backing of the Prime Minister is including different happiness questions in the Integrated Household Survey (a quarterly survey with an annual sample of 200,000), and has outlined methods and measures that can be used to take these data to the forefront of public policy (Dolan, et al., 2011; Fujiwara and Campbell, 2011).

The important thing about happiness data is that they allow us to say what is important in people's lives when they are not thinking about how important those things are. I cannot overstate how important this is. Using regression analysis to explain happiness based on a range of factors that cause happiness allows us to examine the effects of a range of causes on the final consequence -- without asking people to do it for us. We can find out how important walking with a cane is alongside all the other things that affect happiness. This frees us from depending on the cognitive capacity of respondents to simultaneously weigh up different levels of different dimensions in a preference elicitation study. Instead, the constraints are the degrees of freedom in a regression model and the length of a survey about background variables, where these are required from self-reports (as opposed to already being available from medical and other records). All of this means that we can reduce reliance on crude and narrow descriptive systems such as the EQ-5D and SF-6D and not expect them to pick up everything that is important about a health state.

The lack of clarity about what dimensions and levels to use in a health state description arises partly out of confusion over what it is we are ultimately trying to value. The literature demonstrates the common tendency to make distinctions among the often-confused terms of health related quality of life, quality of life and wellbeing (Taillefer, at al., 2003). The World Health Organisation's constitution defines health as "a state of complete physical, mental, and social wellbeing and not merely the absence of disease and infirmity" (WHO, 1948). Many attempts have since been made to either narrow this definition (Cummins, et al., 2004), or to make it so broad as to include all possible economic and societal influences (Shilton, et al., 2011). The distinction is very blurred, however, and it actually misses the point. Happiness should be seen as the final consequence of policy intervention in any area of public policy – the left-hand side or the dependent variable in a regression analysis. We may quite legitimately decide that particular determinants of happiness – the right-hand side or independent variables – are the main responsibility of different government departments (health care focuses on affecting health-related determinants), but we should not seek to decompose the overarching objective of policy, not least because something quite significant may end up falling into the gaps.

Constructing descriptive systems often clouds the central issue of these endeavours, which should be to accurately capture the effects of treatments on people's lives. Even within the generic measures, differences exist about what is in and what is out of a measure of health. The Health Utilities Index (HUI), for example, adopts a narrower "within-the-skin" definition of health as compared to the broader definition found in the dimensions of the EQ-5D and SF-6D (Furlong, et al., 2001). Current moves to increase the sensitivity of the EQ-5D by adding further levels to the existing dimensions arguably would be better directed at accounting for missing dimensions, such as vitality. The inclusion of a vitality domain could lead to the EQ-5D+V (or other types of "bolt-on" [EuroQol, 2011]), but this still requires commitment to a rigid descriptive system across all conditions.

Health state descriptive systems have a difficult balance to strike between the need for sensitivity (with all attributes that matter assessed at the appropriate number of levels for discrimination) and the need for a limited number of dimensions and levels to allow for valuation. With too few levels and dimensions, the system lacks sensitivity; with too many levels, the valuation exercise becomes cumbersome and unwieldy due to the large number of health states involved. An

approach that uses happiness as the measure of health care benefit allows us to consider whether we actually need to pre-select specific dimensions of health as having privileged status. It would allow us to examine the impact of health care (and other domains of life) on an individual's happiness using a regression analysis, where information regarding happiness is collected alongside health information and more general background information.

Using this approach, gains (and losses) of happiness attributable to individual health technologies can be derived from both focused clinical trials and from larger scale population studies. This may go some way towards dealing with the problems of the lack of sensitivity and coverage of the existing "right-hand side" health state descriptions and simultaneously address concerns with preference utility scales on the "left-hand side" of a regression model. Returning to the questions of what is to be valued, how it is to be valued and who is to value it, an approach using these methods would allow for a descriptive if so required (the what), but what would elicit happiness ratings directly (the how) in all those affected by an intervention or policy (the who). If we accept that there may be a case for valuing health directly, and that there is an argument for reconsidering the use of descriptive systems, then we need fuller and more considered discussion of how best to do this.

All of this raises serious concerns for the developers of new health technologies and for NICE – at least insofar as they are concerned with valuing what matters most to those affected by health conditions. Of course, no "gold standard" exists against which to determine what matters most, but we can ground the discussion in a simple question: What matters to people when their happiness is explained in terms of their health? When we look at the determinants of happiness, we find that health is one of the most important dimensions; as Graham (2008) points out, "health is among the handful of measurable variables that account for observed variability in human happiness". In the only paper we could find that elicited EQ-5D and happiness, Graham, et al. (2009) showed that anxiety/depression was strongly negatively correlated with life satisfaction in a South American population; in contrast, mobility was much more weakly correlated.

Beyond this, the dimensions that matter most have not been considered closely in the literature thus far; the emphasis has been on broad categories of disability (Oswald and Powdthavee, 2008) or specific health conditions, such as stroke and acute myocardial infarction (Shields and Wheatley Price, 2005) that have been shown to reduce happiness. Having said that, and unsurprisingly, a strong negative association exists between happiness and mental health (Diener and Seligman, 2002). The penetration of mental health conditions into our lives and on our happiness must be due, in part at least, to difficulties in adapting to these conditions compared to other diagnoses. We know, for example, that people with chronic health conditions or who are physically disabled show considerable levels of adaptation to these conditions (Hurst, et al., 1994). The improvement in happiness (although not to pre-morbidity levels) has been explained by the fact that paraplegia and the like are "part-time" experiences that affect wellbeing only when attention is drawn to the limitations (Dolan and Kahneman, 2008). This is not to trivialise such conditions, but to more accurately place them in the context of the richness of our lives. In contrast, many mental health problems are more "full-time" in their attention-seeking and impact on our lives.

Happiness also is affected by less-attention-seeking conditions. In an excellent study using data on 16 countries, Blanchflower and Oswald (2008) show that nations with higher happiness also report systematically lower levels of hypertension. This is interesting because hypertension is often asymptomatic and so there may also be latent effects of a condition on happiness. It is generally held that our happiness is improved by effective medical treatment and numerous studies demonstrate that happiness improves when health care interventions reduce symptoms (e.g., Pressman and Cohen, 2005). Results of this kind cast some doubt on the popular idea that we have an in-built "baseline" level of happiness to which we return. The paraplegics in the classic Brickman, et al. (1978) study were still less happy after a year than the control group, just not as miserable as those focusing on their condition would expect them to be. A genetic predisposition to have different baselines (De Neve, et al., 2010) may exist, but some adapting to some conditions, such as depression, simply does not happen. Given these findings, I am increasingly concerned that mental health may be undervalued using existing measures (Dolan, et al., 2009, Dolan, et al., 2011).

4. HAPPINESS-BASED TARIFFS FOR THE SF-6D AND EQ-5D

Key points:

- In our analyses of the SF-6D and the EQ-5D, mental health appears most important to happiness, with vitality also important in the SF-6D
- Physical functioning is not significantly negatively associated with happiness
- The effect that pain has on happiness is not consistent and requires further investigation
- Preference-based tariffs show that physical functioning and pain matter more to people as compared to the small effects these dimensions have on assessments of happiness

4.1. Data from the British Household Panel Survey

Since 1991, the BHPS has annually surveyed each adult member of a nationally representative sample of about 5,500 households, reaching a total of approximately 10,000 individual interviews (rising to about 15,500 by 2005). The survey now is called "Understanding Society" and the sample sizes have roughly quadrupled. From 1997 (excluding 2001) respondents have been asked "how dissatisfied or satisfied are you with your life overall" using a response scale from 1 (not satisfied) to 7 (completely satisfied). The responses are rescaled onto a 0-1 scale (1 maps to 0 and 7 maps to 1). The average rescaled score is 0.705 (the standard deviation is 0.217). The SF-36 was included in 1999 and 2004 (waves 9 and 14), thereby allowing happiness-based weights for the SF-6D to be estimated in those years. Table 1 summarises the distribution of SF-6D responses in the BHPS.

Table 1: Distribution of SF-6D responses in the BHPS

	Physical functioning	Role limitations	Social functioning	Pain	Mental health	Vitality
Level 1	53.8%	76.9%	72.6%	47.0%	41.4%	5.4%
Level 2	26.8%	10.9%	9.7%	21.4%	30.2%	44.8%
Level 3	9.3%	4.6%	9.3%	14.7%	22.0%	29.6%
Level 4	3.6%	7.5%	6.1%	6.6%	4.7%	11.7%
Level 5	4.4%		2.3%	7.5%	1.7%	8.4%
Level 6	2.0%			2.7%		

The impact on happiness of each level for each dimension of the SF-6D may be estimated through a linear model predicting life satisfaction using the SF-6D levels as determinants. Since the SF-6D can be calculated for two waves of data, we have a panel that may be analysed using either fixed effects (FE) or random effects (RE). FE estimates the coefficients from a "within" person comparison, comparing each individual to him/herself at a different point in time when circumstances are likely to be different (effectively removing the unobserved individual effect from the model). RE estimates the coefficients from an average of the "within" person comparison and a "between" person, cross-sectional comparison (treating the unobserved individual effect as a random variable with a mean of 0). RE utilises more information, making the estimators more efficient (Verbeek, 2000), but it will result in bias and inconsistent estimators if the unobserved individual effect is correlated with any of the independent variables (Ferrer-i-Carbonell and Frijters, 2004).

The impact of the SF-6D is presented as a reduced model of only dimensions/levels, but also including a number of other control variables. These include gender, age, age-squared, gross household income, marital status, employment status, and a year dummy to control for fixed time effects. The impact of the SF-6D on life satisfaction is shown in Table 2. The relatively good explanatory power of the models underlines the importance of health compared to other factors that affect life satisfaction. The addition of control variables has minimal impact upon the size of the coefficients, and so for simplicity we will focus only on the reduced model. The relatively good explanatory power of the models underlines the importance of health compared to other factors that affect life satisfaction. The Hausman test (Hausman, 1978) rejects the hypothesis of no systematic differences between the coefficients from the random effects and fixed effects models suggesting the superiority of the fixed effects model. As expected, in the fixed effects analysis, when looking at only "within" person change, the limited variation across variables between the two periods results in increased standard errors and consequently fewer significant coefficients.

When either the control variables (3) are added or individual fixed effects (2) are controlled for, the impact on life satisfaction falls into line with expectations. When only age is added as a control, the coefficients also are similar to the addition of the full set of controls, suggesting this effect is caused by age. Age is negatively correlated to physical functioning, but shows a non-linear relationship to life satisfaction such that the decline in physical functioning occurring in old age is compensated for by an age-related increase in life satisfaction. Mental health provides the largest detriment of all the dimensions. Interestingly, the impact is roughly halved when the unobserved individual effects are controlled for in the fixed effects models. This suggests that some of the reduction in life satisfaction may be caused by the individual effects, such as personality. Vitality follows a similar pattern to mental health. The fact that the coefficients are not monotonic for the physical and role dimensions is not easily explained at this stage. It is likely that the functional link between the physical and role and social functioning domains and happiness will go some way to explaining this and future focused work will help unpick it.

Table 2: Impact of the SF-6D on life satisfaction (scaled as 0 to 1)

	(1) RE – reduced	(2) FE – reduced	(3) RE – Full	(4) FE - Full
Physical 2	0.017*** (0.003)	0.003 (0.005)	-0.004 (0.003)	0.003 (0.006)
Physical 3	0.022*** (0.005)	-0.009 (0.007)	-0.010** (0.005)	-0.006 (0.010)
Physical 4	0.033*** (0.007)	-0.018 (0.012)	-0.009 (0.009)	-0.009 (0.018)
Physical 5	0.014** (0.007)	-0.028** (0.012)	-0.028*** (0.008)	-0.019 (0.018)
Physical 6	-0.008 (0.010)	-0.091*** (0.019)	-0.054*** (0.014)	-0.072** (0.034)
Role 2	0.011** (0.004)	0.005 (0.006)	0.007 (0.005)	0.005 (0.009)

	(1) RE – reduced	(2) FE – reduced	(3) RE – Full	(4) FE - Full
Role 3	-0.065*** (0.006)	-0.045*** (0.010)	-0.061*** (0.007)	-0.044*** (0.014)
Role 4	-0.043*** (0.006)	-0.038*** (0.009)	-0.041*** (0.007)	-0.039*** (0.013)
Social 2a	-0.021*** (0.004)	-0.010* (0.006)	-0.019*** (0.004)	-0.009 (0.007)
Social 3	-0.032*** (0.004)	-0.023*** (0.007)	-0.029*** (0.005)	-0.021** (0.010)
Social 4	-0.069*** (0.006)	-0.054*** (0.009)	-0.062*** (0.008)	-0.050*** (0.014)
Social 5	-0.108*** (0.009)	-0.043*** (0.016)	-0.105*** (0.013)	-0.032 (0.027)
Pain 2	-0.007** (0.003)	-0.001 (0.004)	-0.010*** (0.003)	-0.001 (0.005)
Pain 3	-0.012*** (0.004)	-0.006 (0.005)	-0.012*** (0.003)	-0.005 (0.007)
Pain 4	-0.022*** (0.005)	-0.007 (0.008)	-0.020*** (0.006)	-0.008 (0.011)
Pain 5	-0.029*** (0.006)	-0.013 (0.009)	-0.021*** (0.007)	-0.013 (0.013)
Pain 6	-0.045*** (0.009)	-0.048*** (0.014)	-0.029** (0.011)	-0.042* (0.022)
Mental 2	-0.062*** (0.003)	-0.037*** (0.004)	-0.054*** (0.002)	-0.036*** (0.005)
Mental 3	-0.106*** (0.003)	-0.066*** (0.005)	-0.097*** (0.003)	-0.066*** (0.007)
Mental 4	-0.162*** (0.006)	-0.090*** (0.010)	-0.143*** (0.008)	-0.087*** (0.016)
Mental 5	-0.214*** (0.009)	-0.145*** (0.016)	-0.190*** (0.013)	-0.143*** (0.029)
Vitality 2	-0.046*** (0.005)	-0.020** (0.008)	-0.050*** (0.005)	-0.019* (0.010)
Vitality 3	-0.090*** (0.005)	-0.044*** (0.009)	-0.096*** (0.005)	-0.042*** (0.011)
Vitality 4	-0.140*** (0.007)	-0.085*** (0.011)	-0.147*** (0.007)	-0.084*** (0.015)
Vitality 5	-0.155*** (0.007)	-0.097*** (0.012)	-0.168*** (0.008)	-0.093*** (0.018)

	(1) RE – reduced	(2) FE – reduced	(3) RE – Full	(4) FE - Full
Dummy to pick up whether dimension is at the worst level	0.026*** (0.005)	0.014* (0.007)	0.023*** (0.005)	0.012 (0.010)
Log hh income			0.008*** (0.002)	0.005 (0.004)
Age 21-25			-0.016*** (0.006)	0.017 (0.033)
Age 26-30			-0.017*** (0.006)	0.047 (0.063)
Age 31-35			-0.028*** (0.006)	0.067 (0.094)
Age 36-40			-0.034*** (0.006)	0.090 (0.124)
Age 41-45			-0.037*** (0.007)	0.104 (0.154)
Age 46-50			-0.030*** (0.007)	0.119 (0.184)
Age 51-55			-0.011 (0.007)	0.149 (0.215)
Age 56-60			0.018** (0.007)	0.197 (0.245)
Age 61-65			0.039*** (0.008)	0.238 (0.276)
Age 65-70			0.067*** (0.009)	0.269 (0.306)
Age 71-75			0.083*** (0.010)	0.293 (0.336)
Age 76-80			0.086*** (0.011)	0.302 (0.367)
Age 81-98			0.094*** (0.012)	0.318 (0.397)
Male			-0.032*** (0.003)	0.000 (0.000)
Unemployed			-0.039*** (0.008)	-0.030* (0.018)
Long term sick			-0.045*** (0.008)	-0.067*** (0.021)
Retired			0.002 (0.006)	-0.009 (0.012)

	(1) RE – reduced	(2) FE – reduced	(3) RE – Full	(4) FE - Full
Family carer			-0.007 (0.005)	-0.018 (0.013)
Job other			0.009* (0.005)	0.015 (0.012)
Married			0.027*** (0.003)	0.008 (0.010)
Separated			-0.047*** (0.009)	-0.035* (0.020)
Divorced			-0.037*** (0.006)	-0.001 (0.019)
Widowed			0.001 (0.007)	-0.019 (0.027)
Wave nine			0.003 (0.002)	0.029 (0.030)
Constant	0.846*** (0.005)	0.796*** (0.008)	0.802*** (0.015)	0.615*** (0.187)
Observations	27865	27865	27773	27773
R-squared	0.216	0.093	0.264	0.102
Number of individuals	19230	19230	19185	19185

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

4.2. New data from the US

Our new dataset elicited responses for the EQ-5D and two measures of happiness: life satisfaction and day affect. The survey began with a life satisfaction question: *"Overall how satisfied are you with your life?" (on a 0-6 scale) followed by questions relating to affect yesterday: "Overall, how did you feel yesterday? Please rate each feeling on the scale given. A 0 means that you did not experience that feeling at all and a 6 means that you experienced that feeling very strongly".* The feelings were: friendly, lethargic, stressed, happy, sad, calm, angry, tired, depressed, and worried. "Day affect" is taken as the difference between the average of positive and negative affect and is highly discriminatory in terms of an individual's moment-to-moment experiences. This proxies the influential day reconstruction method developed by Kahneman (Kahneman, et al., 2004) that divides an individual's experiences over the previous day into discrete activities that are then rated according to the mood associated with them. All measures are rescaled onto a 0-1 scale.

The questionnaire was administered via on-line and telephone interviews with respondents who are part of a panel of members of the general public that is used by RAND in various health surveys. Many issues about selection bias in relation to Internet and telephone samples exist, but this study was not designed to elicit responses from a truly representative sample of the population; instead, it was meant to advance understanding of the relationship between happiness and health. All 1173 respondents are included in the analysis. The sample had an average age of 50, 46% were male, 65% were married and 62% were in paid employment. Average household income was $54,237. Table 3 shows the distribution of EQ-5D in the sample.

Table 3: US sample – EQ-5D

	Mobility	Self-care	Usual activities	Pain	Mood
Level 1	81.1%	96.9%	82.7%	50.7%	67.1%
Level 2	18.6%	2.9%	16.5%	46.6%	30.9%
Level 3	0.3%	0.3%	0.8%	2.6%	2.1%

Table 4 presents two specifications for each dependent variable. The restricted specification does not control for any background variables whilst the full specification (as for the BHPS) controls for income, age, gender, employment, marital status and race. For life satisfaction, the only domains that remain significant for the full specification are (in order of importance) mood, usual activities, and pain. For day affect, mood also matters the most, followed again by pain and then usual activities. The relative effect of mood is much greater, which might be expected for a more experience-based measure of happiness.

Table 4: Impact of EQ-5D on happiness

	Life satisfaction	Life satisfaction	Day affect	Day affect
EQ1 - Mobility	0.000 (0.016)	-0.003 (0.016)	0.017 (0.014)	-0.001 (0.014)
EQ2 – Self-care	-0.065* (0.037)	-0.016 (0.037)	-0.051** (0.024)	-0.033 (0.025)
EQ3 – Usual act.	-0.052*** (0.018)	-0.038** (0.017)	-0.026* (0.015)	-0.027* (0.015)
EQ4 – Pain	-0.023** (0.010)	-0.022** (0.010)	-0.042*** (0.009)	-0.050*** (0.009)
EQ5 – Mood	-0.160*** (0.011)	-0.142*** (0.011)	-0.200*** (0.009)	-0.185*** (0.010)
Ln(HH income)		0.054*** (0.010)		-0.006 0.008
Male		-0.033*** (0.009)		-0.016 0.009
Age		-0.004* (0.002)		0.004* 0.002

	Life satisfaction	Life satisfaction	Day affect	Day affect
Age2		0.000** (0.000)		0.000** 0.000
Unemployed		-0.058** (0.030)		-0.014 0.028
Sick		-0.119 (0.084)		-0.081* 0.048
Disabled		-0.102*** (0.035)		-0.038 0.029
Retired		0.028* (0.015)		0.018 0.014
Homemakers		0.035** (0.017)		0.009 0.019
Job other		0.051* (0.028)		0.008 0.023
Married		0.030* (0.015)		0.001 0.014
Separated		-0.007 (0.054)		0.007 0.044
Divorced		-0.014 (0.019)		-0.009 0.017
Widowed		0.002 (0.024)		-0.012 0.024
Race - black		-0.006 (0.020)		-0.008 0.021
Race - Indian		-0.011 (0.059)		-0.012 0.057
Race – Asian		-0.072** (0.034)		-0.025 0.028
Race – other		-0.011 (0.044)		-0.016 0.046
R2	0.29	0.37	0.38	0.44
Obs	1173	1169	1170	1166

Robust standard errors used. ***, **, * represents significance at the 1%, 5% and 10% levels respectively.

Table 5 uses the full specification for each measure of happiness and treats the EQ-5D in the same ways as Dolan (1997). This method sets out the coefficients on each of the levels and dimensions in the EQ-5D so that overall valuations for composite health states can be derived.

For life satisfaction, extreme anxiety/depression again has the largest negative association with life satisfaction. This is followed closely by being unable to perform usual activities and moderate anxiety/depression. Moderate pain has a negative association with life satisfaction, but this impact is much lower than the other negative variable and it is important to note that not all levels of pain have a negative association with life satisfaction. The impact of usual activities and anxiety/depression are much larger than other variables, such as unemployment and income, that are found in the literature (and indeed in these equations). For day affect, extreme and moderate anxiety/depression have the largest impact. The next largest domain is pain, where both extreme and moderate pain have large negative associations with day affect.

Table 5: Levels of the EQ-5D and happiness

EQ-5D domain	Life Satisfaction	Day Affect
Mobility 2	-0.016 0.016	-0.002 0.015
Mobility 3	0.185* 0.111	0.013 0.068
Self care 2	0.021 0.032	-0.033 0.026
Self care 3	-0.120 0.141	-0.067 0.101
Usual activities 2	-0.018 0.017	-0.026 0.016
Usual activities 3	-0.213*** 0.078	-0.030 0.065
Pain/discomfort 2	-0.021** 0.010	-0.047*** 0.010
Pain/discomfort 3	-0.047 0.075	-0.117* 0.063
Anxiety/depression 2	-0.134*** 0.011	-0.182*** 0.011
Anxiety/depression 3	-0.280*** 0.076	-0.388*** 0.060
Any dimension at level 3	-0.053 0.076	-0.005 0.059
Obs	1169	1166
R2	0.38	0.44

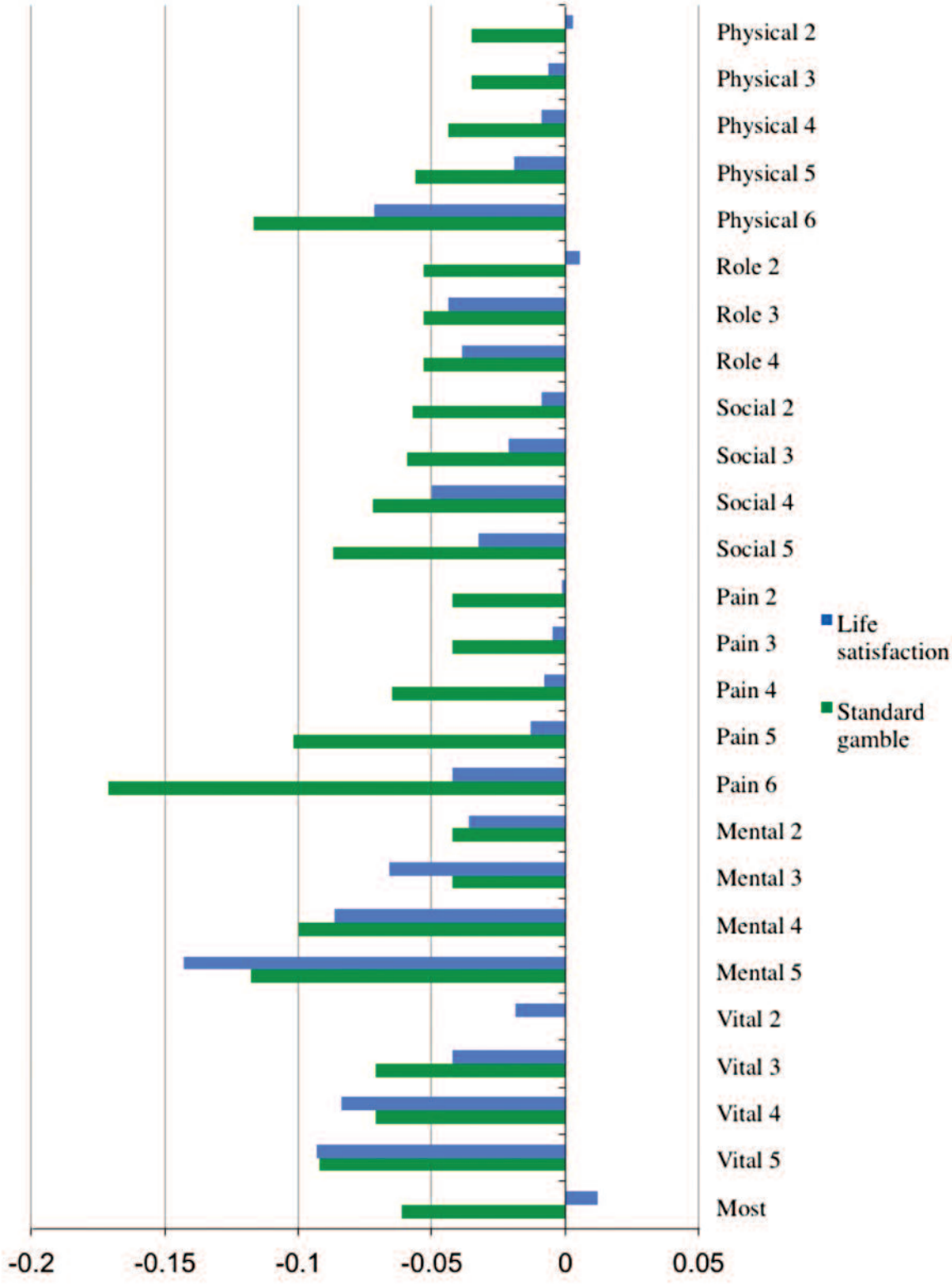
Notes: Controlling for the background variables in the previous tables. Robust standard errors used. ***, **, * represents significance at the 1%, 5% and 10% levels respectively.

4.3. Comparing the tariffs

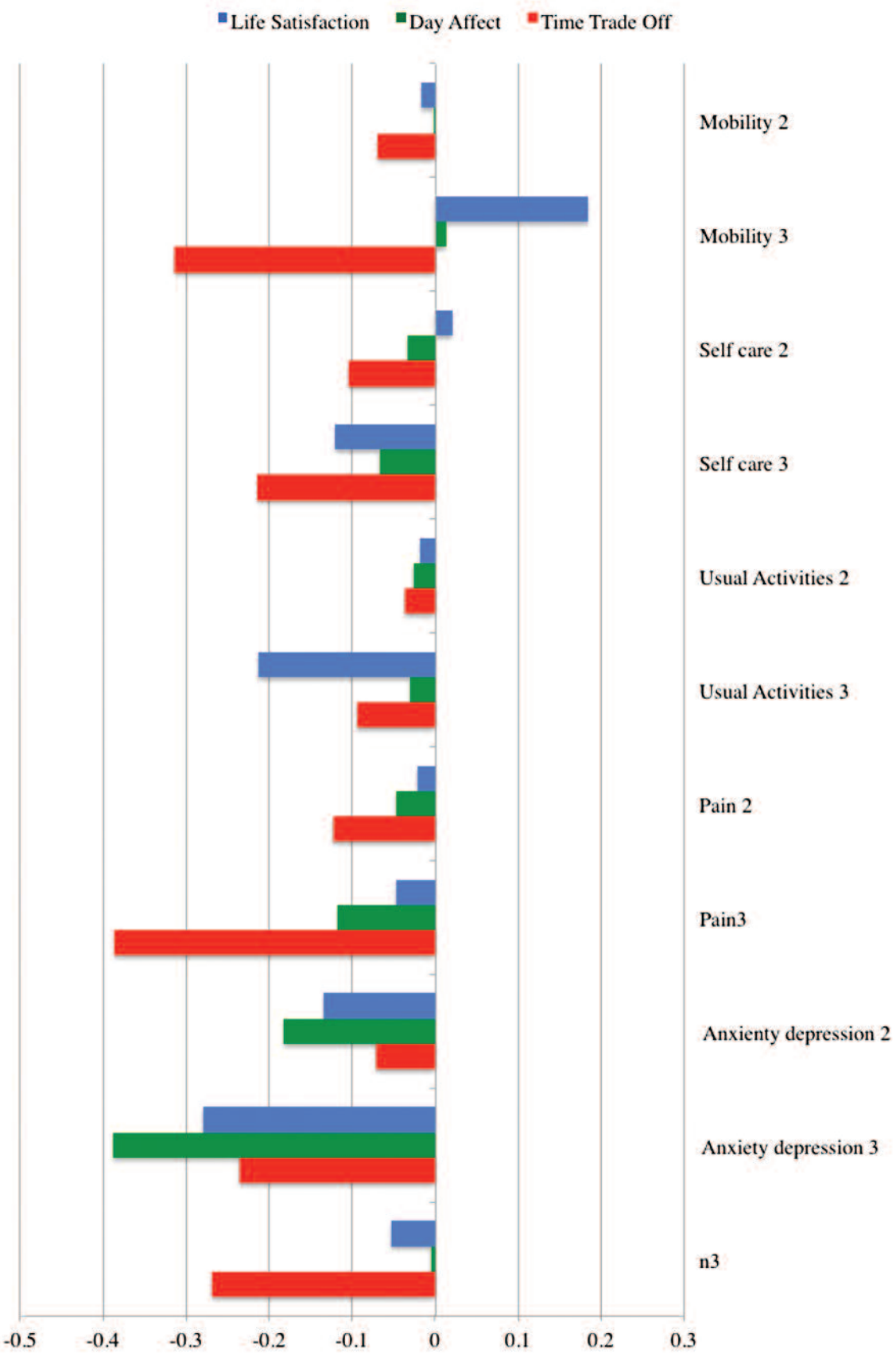
Graph 1 compares the preference-based weights estimated from standard gamble responses with the happiness-based weights for the SF-6D estimated from life satisfaction reports in the BHPS (FE-reduced model). Physical functioning, role functioning, social functioning and pain all show up as more important in SG responses than in happiness regressions. Mental health and vitality have about the same relative impact on preferences and happiness. These results mean that most health states would be more severe using preferences compared to happiness, but that mental health and vitality would loom relatively large using happiness.

Graph 2 compares the preference-based weights estimated from TTO responses with the happiness-based weights for the EQ-5D estimated from life satisfaction and day affect reports in the US data. Self care, mobility and pain all show up as more important in TTO responses compared to the happiness regressions. The reverse is true for anxiety/depression, where the impact is greater on happiness. Looking at differences between happiness measures, mental health and pain show up as more important for day affect than for life satisfaction for both levels of severity. In the most severe levels, usual activities and self care show up as more important for life satisfaction, but at moderate severity these dimensions matter more for day affect. Notice the positive coefficient on "confined to bed". This is likely to be due to small samples. Overall, looking generally at both graphs, it would appear that pain matters much more in the imagination than in the experience and that mental health matters relatively more in the experience than in the imagination.

Graph 1: Comparison of happiness and preference-based weights for the SF-6D



Graph 2: Comparison of happiness and preference-based weights for the EQ-5D



5. DISCUSSION

- Adopting happiness as a measure for use in health technology assessment allows a more direct way of valuing the impact of health states on the real experiences of all those affected by interventions at all stages in the process
- The happiness-based tariffs reported here can be used to see what difference they would make to the cost-per-QALY of existing and new technologies
- Future research should seek to gather more data on happiness and to better understand the source of the difference between preference-based and happiness-based valuations

As noted in Chapter 3, many muddled attempts have been made to draw boundaries around what constitutes "health-related quality of life". Why bother? I accept there might be some merit in focusing health care interventions on health-related causes of losses in happiness, but the consequences of those causes, like any other cause, are felt – quite literally – by their impact on happiness. It is not at all clear that mobility would be in the "top five" health-related determinants of happiness as it is in the EQ-5D. Vitality would be in there, too, probably alongside sexual functioning and communication (Bowling, 1995). Mental health would undoubtedly matter most of all. Importantly though, we do not have to pick a limited number of dimensions for valuation because happiness data are not constrained by the cognitive capacity of respondents to simultaneously weigh up dimensions in a preference elicitation task.

Even if we were to stick to those dimensions of health prescribed by the EQ-5D and SF-6D, Chapter 4 shows that the dimensions of health that matter most to preferences do not appear to be those that most affect happiness. From our analysis of existing and new data, mental health is found to have the largest and most significant effect on happiness. Whilst mental health matters in the preferences of the general public, its importance relative to physical functioning and pain is underestimated in the imaginations of people. The fact that mental health conditions are among the most resistant to adaptation underlines the longevity of suffering associated with these conditions. The findings of the study by Levack, et al. (2004) on cord compression seems directly relevant here: They showed that the quality of life differences between participants who were able to walk and those not able to walk was small, but psychological issues (as measured by the Hospital Anxiety Depression Scale) affected individuals much more.

Given the importance of mental health as a determinant of happiness, a focus on happiness means that treatments that bring about reductions in anxiety or depression will be shown to be relatively more cost-effective than when using preference-based methods. This still means that treatments for physical functioning, pain, and the like, can be effective – but that they will be even more so if they also have positive consequences for mental health. To restate my clear normative position: The impact upon happiness is the final consequence in the aetiological model. The impact of physical functioning on happiness could then be seen indirectly through its effect on mental health and directly where there are no effects on mental health but happiness is nonetheless lower.

Physical functioning problems may become less severe over time but, at the same time, some of the problems may be "transferred" onto other dimensions, such as vitality. Future studies and analyses should consider how best to capture the direct and indirect effects of a health problem. Ideally, this will involve the collection of health state and happiness data from the same individuals, with more frequent assessments than annually, as for most existing datasets. If, for

example, the impact of physical functioning is through vitality, it makes even less sense to ask people to predict how much these two dimensions will affect them because they are unlikely to forecast the complex interaction effects, especially as these change over time. SG and TTO will effectively mix up the direct and indirect effects of the right-hand side variables (even if the left-hand were to be specified appropriately).

In common with many new approaches to old problems, the happiness approach is not without criticism. It raises three main concerns: adaptation; anchoring (around death); and measurement.

Adaptation

Issues of adaptation and the changing dynamics of how dimensions interact with one another apply to the description as well as the valuation of health. For example, an individual who previously described him/herself as having "some problems walking about" due to mild arthritis before undergoing a bilateral lower limb amputation, also may describe him/herself as having "some problems walking about" after they have adapted to their condition and begin walking again with prosthetic limbs. In a related way, there could be "reference shifts" within the assessment of each dimension (Carver and Scheier, 2000). So individuals might think that their pain is moderate, but when they then experience extreme pain, it shifts their reference to moderate pain, which might now actually feel like very little pain (Dar, et al., 1995).

This raises a criticism I am sure to face by nailing myself so firmly to the mast of happiness: People adapt to poor health and, all else equal, the more a patient adapts to her condition, the less priority she will receive in the competition for resources that improve quality of life. Sen (1992) has raised similar concerns about a preference satisfaction account of welfare, and it is perhaps why many of those who still favour a preference account cling to the idea that preferences should be elicited from the general public before any adaptation has taken place.

I suggest that this discussion has suffered from a focusing effect of its own: It has focused on the injustice of giving less priority to those who adapt, without properly considering the injustice of giving less priority to those who do not adapt. Giving some people greater priority than is warranted by the size of the actual happiness loss, by definition, results in less priority being given the real losses in happiness experienced by others. Discussions about adaptation should be informed by the specific elements of the adaptation process (Menzel, et al., 2002) but, in general, it is morally suspect to prioritise adaptive conditions or patients when it comes at the expense of non-adaptive ones and necessarily, therefore, when those who are experiencing less suffering are given priority over those who are suffering more.

Forecasts of future utility should therefore take due account of adaptation but, as noted in Chapter 2, the evidence suggests that the intuitive forecasts of lay people generally fail to do so (Gilbert and Wilson, 2000). The results from health state valuation studies are consistent with the idea that people underestimate the extent to which they and others will adapt to changed circumstances. We may find it challenging to imagine or believe the adaptation that appears to take place with some health conditions – but this probably has more to do with the inaccuracy of our imaginations and beliefs than with the legitimacy of the responses of those who have adapted to changed circumstances. The focusing illusion, for example, means that we focus on a salient aspect of an event when, in fact, that aspect has much less effect on us than we predict. As

Schkade and Kahneman (1998) note “nothing in life is quite as important as you think it is while you are thinking about it”. It is almost impossible for any of us (including economists and philosophers) to avoid the focusing illusion.

When the general public and “patients” are asked to imagine what it would be like to be in another health state, they have their attention drawn away from the possibility of adaptation and instead towards transitional changes in “health-related” domains of life defined by the particular health state descriptive system. As Tessa Peasgood pointed out to me about 15 years ago, when I thought that SG and TTO preferences expressed something meaningful about utility, "when you die, you give up everything in life and not just your EuroQol health state". I still think that SG and TTO reflect something meaningful about the affective reaction to a health state: The trouble is that they are not supposed to be based on the shock and fear associated with a health state (which can be better captured by happiness data anyway).

Anchoring

Considerations of death raise a more practical concern about happiness data: Responses are not calibrated on a scale anchored at zero for death, and so they cannot be used to calculate QALYs in the conventional sense. One way round the problem of valuing death (kindly suggested by one of the referees) is to value the probability of death. This is in keeping with the approach taken in most of the value-of-life literature and has the attraction of having a utility interpretation. In principle, life satisfaction will be affected by the probability of death. So, if it is possible to devise questions that measure the probabilities of death, then we could derive a value of these probabilities in terms of their effects on life satisfaction. Even if life satisfaction were to be affected by risk of death (and it may not be), the problem is that it will be affected by perceptions of the risk of death, which often bear little relation to the actual risk of death[2] and some differences by background characteristics are salient (Hakes, et al., 2004)). The estimates from such an approach therefore could be quite biased.

In a more direct way, we could explore anchoring the life satisfaction question on “dead at zero”. Ideally, we would get some comparable within- and across-respondent data on the differences between the widely used endpoints and the endpoint of death. In this way, we might be able to map one set of values to the other, thus adjusting existing data “as if” it were on a scale more comparable to that used in the QALY approach.

In the meantime, I am not at all convinced that we need to bother rating happiness against death at all. Since almost everyone would not rate their current state as worse than dead by any measure of experience, we can estimate the loss in happiness from premature death to be the happiness that would have been expected otherwise. I appreciate that this assumes that the lowest point on existing happiness would be assumed to be equivalent to death, but the fraction of respondents giving the lowest score is very low (for the two waves of the BHPS, 1.5% of responses are 1 out of 7). My suggestion also does not allow for states rated as worse than dead, but the conceptual and practical bases for such states is highly controversial and I suspect, at least from a policy point of view, that nothing much would be lost by ignoring them. We could then estimate QALYs as the area under a happiness curve; not that far from Edgeworth’s (1881) original conception of a hedonimeter. Note also that the top end of the standard QALY scale is truncated at “no problems”, thus failing to capture any benefits at all from feeling happy.

[2] It seems we generally under-estimate the objective risk of death (Anderson, 2011).

Measurement

Happiness measures are not without their problems and I am actively engaged in research that addresses them. They present at least three methodological issues: salience; scaling and selection. On salience, any question focuses attention on something and we must be clear about where we want respondents' attention to be directed, and where it might in fact be directed. We should like to have attention focused on those things that will matter to the respondents when experiencing their lives and thinking about what matters to them, possibly including any thoughts they may have about their health, which might matter more than the state itself (Dolan, 2011). It must be recognised that the mere act of asking a happiness question might affect experiences (Wilson, et al., 1993). Responses will be influenced by salient cues, such as the previous question (Schwarz, et al., 1987), and perhaps also by the organisation carrying out the survey. The general consensus, however, is that there are stable and reliable patterns in happiness, even over the course of many years (Fujita and Diener, 2005).

On scaling, in order to make meaningful comparisons over time and across people, we need to understand how interpretations of the scales may change over time. Frick, et al. (2006) show that respondents in the German Socio-Economic Panel have a tendency to move away from the endpoints over time. The relationship between earlier and later responses can be seen as an issue of scaling and salience: If later responses are influenced by earlier ones, then the earlier ones are salient at the time of the later assessment (as shown in the study by Dolan and Metcalfe, 2010). It is possible that the endpoints on a scale change when circumstances change and when key life events happen; it is important that we conduct more focused empirical research into this issue. It is not at all clear, though, whether this actually matters for policy purposes since a seven out of ten before and after having children, for example, is still, in fact, seven out of ten.

On selection, who chooses to be in a survey with happiness measures is important to establishing whether the effects of any factor associated with happiness are generalisable or specific to the sample population. Attrition of certain types of people in different types of happiness surveys also is important in generalising treatment effects. Watson and Wooden (2004) show that people with lower life satisfaction are less likely to be involved in longitudinal surveys. Moreover, people self-select into particular circumstances that make it difficult for us to say anything meaningful about how those circumstances would affect other people. Take the effects of volunteering as an example. Volunteering and happiness are generally associated positively, but it is possible that those choosing to volunteer are those most likely to benefit from it and those with greater SWB may be those most likely to volunteer in the first place. Part of any correlation will then be picking up the causality from SWB to volunteering. Telling the chicken from the egg in happiness research, as elsewhere, is crucial for effective policy making.

These are challenges, but not fundamental flaws, with happiness. In taking things forward to generate more and better data, happiness needs to be routinely gathered in trials and surveys, ideally from the same individuals over relatively short time periods, and from relatives and carers as well as patients. We must be alert to the fact that in some areas and populations we do not yet know enough about the determinants of happiness, their relative coefficients and how endogenous and exogenous factors influence these data to make bold policy decisions. In practical terms, the best way to address this is by incorporating happiness measures into mandatory returned data under existing programmes within the NHS. For example, the

Department of Health's contract for Acute Services stipulates that patient reported outcome measures (as well as the EQ-5D) are returned for key surgical conditions pre- and post-treatment. Incorporating happiness measures into such initiatives at a regional and national level would provide data to shed further light on the dynamic nature of happiness and its determinants.

We do not need to wait for more research before making some tentative recommendations. As things stand, we can estimate SF-6D and EQ-5D tariffs based on the impact these measures have on happiness. The most robust tariff for the SF-6D would be the one generated using life satisfaction from the BHPS. We do not currently have such good UK data from which to estimate a tariff for the EQ-5D, but we can begin by using the tariff for life satisfaction from the US data. The obvious next step would be to see what difference using happiness-based tariffs would make to the cost-per-QALY of existing or new therapies as compared to using the established preference-based tariffs.

If we are interested in happiness as the final consequence (and there is some support for this from public opinion – see Dolan, et al., 2011), then we should seek to establish the main determinants of happiness and focus policy efforts on them. This paper has suggested some of the implications that this focus would have when valuing different dimensions of health for use in determining priorities in health care, implicitly assuming a fixed budget for health care. It is for future research to consider what implications a focus on happiness would have for decisions about the total expenditure on health care compared to other uses of public resources (and indeed for decisions about the total level of public spending itself).

I do not doubt that preferences should be accounted for when allocating resources and I also do not doubt that opportunities, capabilities and "what people can do" all matter (Hausman and McPherson, 2009). But all these things only matter because they show up in better experiences – maybe not today, maybe not tomorrow, but at some point for someone, somewhere. As a "happiness economist", I make no great claims for the significance of anything – being able to walk or having children – beyond its effect on happiness. As a more recent "happiness parent", I want my children to pursue projects that make them happy. To paraphrase Roger Crisp: "happiness seems just obviously worth having in life, and that is the end of it."

REFERENCES

Anderson, H. (2011) Perception of own death risk: An assessment of road-traffic mortality risk. Risk Analysis. 31(7), 1069-1082.

Bentham, J. (1789) An introduction to the principles of morals and legislations. Oxford University Press.

Blanchflower, D. and Oswald, A. (2008) Hypertension and happiness across nations. Journal of Health Economics. 27(2), 218-223.

Bowling, A. (1995) What things are important in people's lives? A survey of the public's judgments to inform scales of health related quality of life. Social Science and Medicine. 41(10), 1447-1462.

Brazier, J.E., Roberts, J.R. and Deverill, M. (2002) The estimation of a preference-based measure of health from the SF36. Journal of Health Economics. 21(2), 271-292.

Brazier, J.E., Rowen, D. and Hanmer, J. (2008) Revised SF-6D scoring programmes: A summary of improvements. Patient Reported Outcomes Newsletter. (40), 14-15.

Brickman, P., Coates, D. and Janoff-Bulman, R. (1978) Lottery winners and accident victims: Is happiness relative? Journal of Personality and Social Psychology. 36(8), 917-927.

Calvo, E., Haverstick, K. and Sass, S.A. (2007) What makes retirees happier: A gradual or 'cold turkey' retirement? Paper presented at the 9th Annual Joint Conference of the Retirement Research Consortium. 9-10 August. Washington, D.C.

Carver, C. S. and Scheier, M. F. (2000) Scaling back goals and recalibration of the affect system are processes in normal adaptive self-regulation: Understanding 'response shift' phenomena. Social Science and Medicine. 50(12), 1715-1722.

Chalkidou, K., Tunis, S., Lopert, R., Rochaix, L., Sawicki, P.T., Nasser, M. and Xerri, B. (2009) Comparative effectiveness research and evidence-based health policy: Experience from four countries. Milbank Quarterly. 87(2), 339-367.

Cummins, R., Lau, A. and Stokes, M. (2004) HRQOL and subjective well-being: Non-complementary forms of outcome measurement. Expert review of pharmacoeconomics & outcomes research. 4(4), 413-420.

Dar, R., Ariely, D. and Frenk, H. (1995) The effect of past injury on pain threshold and tolerance. Pain. 60(2), 189-193.

Darzi, A. (2008) High quality care for all: NHS Next Stage Review final report. London: Department of Health.

De Neve, J-E., Christakis, N.A., Fowler, J.H. and Frey, B.S. (2010) Genes, economics, and happiness. Basle: Center for Research in Economics, Management and the Arts.

De Wit, G.A., Busschbach, J.J. and De Charro, F.T. (2000) Sensitivity and perspective in the valuation of health status: Whose values count? Health Economics. 9(2), 109-126.
Department of Health (UK). (2010) A new value-based approach to the pricing of branded medicines – a consultation. Available at http://www.dh.gov.uk/en/Consultations/Liveconsultations/DH_122760 [Accessed 26 September 2011].

Diener, E. (2005) Income and happiness. Observer. 18(4). Available at http://www.psychologicalscience.org/index.php/publications/observer/2005/april-05/income-and-happiness.html [Accessed 26 September 2011].

Diener, E. and Seligman, M. (2002) Very happy people. Psychological Science. 13(1), 81-84.
Dolan P. (2011) Thinking about it: Thoughts about health and valuing QALYs. Health Economics. 20(12), 1407-1416.

Dolan P., Lee, H. and Peasgood, T. (in press) Losing sight of the wood for the trees: Some issues in describing and valuing health and another possible approach. Pharmacoeconomics.
Dolan, P. (1997) Modelling valuations for EuroQol health states. Medical Care. (35)11, 1095-1108.

Dolan, P. (2000) The measurement of health-related quality of life for use in resource allocation decisions in health care. In Culyer, A.J. and Newhouse, J. eds. Handbook of Health Economics. Amsterdam: Elsevier North Holland. 1723-1760.

Dolan, P. and Kahneman, D. (2008) Interpretations of utility and their implications for the valuation of health. Economic Journal. 118(525), 215-234.

Dolan, P. and Metcalfe, R. (2010). 'Oops . . . I did it again': Repeat focusing effects in reports of happiness. Journal of Economic Psychology. 31(4), 732-737.

Dolan, P. and White, M. P. (2007) How can measures of subjective well-being be used to inform public policy? Perspectives on Psychological Science. 2(1), 71-85.

Dolan, P., Gudex, C., Kind, P. and Williams, A. (1996) The time trade-off method: Results from a general population study. Health Economics. 5(2), 141-154.

Dolan, P., Layard, R. and Metcalfe, R. (2011) Measuring subjective well-being for public policy. London: Office for National Statistics.

Dolan, P., Lee, H., King, D. and Metcalfe, R. (2009) Valuing health directly. British Medical Journal. 339, b2577.

Edgeworth, F.Y. (1881) Mathematical psychics: An essay on the application of mathematics to the moral sciences. London: Kegan Paul.

REFERENCES

EuroQol Group (1990) EuroQol--a new facility for the measurement of health-related quality of life. The EuroQol Group. Health Policy. 16(3), 199-208.

EuroQol Group (2011) EuroQol EQ-5D products. Available at http://www.euroqol.org/eq-5d/eq-5d-products.html [Accessed 26 September 2011].

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

Fisher, I. (1918) Is "utility" the most suitable term for the concept it is used to denote? American Economic Review. 8, 335.

Frey, B. and Stutzer, A. (2002) Happiness and economics: How the economy and institutions affect human well-being. Princeton: Princeton University Press.

Fujiwara, D. and Campbell, R. (2011) Valuation techniques for social cost benefit analysis: Stated preference, revealed preference and subjective well-being approaches. A discussion of the current issues. London: HM Treasury.

Furlong, W.J., Feeny, D.H., Torrance, G.W. and Barr, R.D. (2001) The Health Utilities Index (HUI®) system for assessing health-related quality of life in clinical studies. Annals of Medicine. 33(5), 375-384.

Gilbert, D.T. and Wilson, T.D. (2000) Miswanting: Some problems in the forecasting of future affective states. In Forga, J. ed. Feeling and thinking: The role of affect in social cognition. New York: Cambridge University Press. 178–197.

Gold, M., Siegel, J.E., Russell, L.B. and Weinstein, M.C. eds. (1996) Cost-effectiveness in health and medicine. New York: Oxford University Press.

Graham, C. (2008) Happiness and health: Lessons—and questions—for public policy. Health Affairs. (27)1, 72–87.

Graham, C., Jaramillo, L.F.H. and Lora, E.A. (2009) Valuing health conditions: Insights from happiness surveys across countries and cultures. Working Paper 4625. Washington, D.C.: Inter-American Development Bank.

Grieve, R., Grishchenko, M. and Cairns, J. (2009) SF-6D versus EQ-5D: Reasons for differences in utility scores and impact on reported cost-utility. European Journal of Health Economics. 10(1), 15-23.

Hakes, J.K. And Viscusi, W.K. (2004) Dead reckoning: Demographic determinants of the accuracy of mortality risk perceptions. Risk Analysis, 24(3), 651-664.

Harsanyi, J. (1985) Does reason tell us what moral code to follow and, indeed, to follow any moral code at all? Ethics. 96, 42-55.

Hausman, D. and McPherson, M. (2009) Preference satisfaction and welfare economics. Economics and Philosophy. 25(1), 1-25.

Hausman, J. A. (1978) Specification tests in econometrics. Econometrica 46(6), 1251-1271.

HM Treasury (UK). (2003) The Green Book. Available at http://www.hm-treasury.gov.uk/data_greenbook_index.htm [Accessed 26 September 2011].

Hurst, N.P., Jobanputra, M., Hunter, M., Lambert, M., Lockhead, A. and Brown, H. (1994) Validity of EuroQol—a generic health status instrument—in patients with rheumatoid arthritis. Rheumatology. 33(7), 655–662.

Kahneman, D. and Riis, J. (2005) Living, and thinking about it: Two perspectives on life. In: Huppert, F.A., Baylis, N. and Keverne, B. eds. The science of well-being. Oxford: Oxford University Press. 285-304.

Kahneman, D., Krueger A.B., Schkade D., Schwarz, N. and Stone, A. (2004). A survey method for characterizing daily life experience: The day reconstruction method. Science. 306(5703), 1776-1780.

Knabe, A., Ratzel, S., Schob, R. and Weimann, J. (2010) Dissatisfied with life but having a good day: Time-use and well-being of the unemployed. Economic Journal. 120(547), 867-889.

Levack, P., Graham, J. and Kidd, J. (2004) Listen to the patient: Quality of life of patients with recently diagnosed malignant cord compression in relation to their disability. Palliative Medicine. 18(7), 594-601.

Menzel, P., Dolan, P., Richardson, J. and Olsen, J.A. (2002) The role of adaptation to disability and disease in health state valuation: A preliminary normative analysis. Social Science and Medicine. 55(12), 2149-2158.

Oswald, A.J. and Powdthavee, N. (2008) Does happiness adapt? A longitudinal study of disability with implications for economists and judges. Journal of Public Economics. 92(5-6), 1061-1077.

Pressman, S.D. and Cohen, S. (2005) Does positive affect influence health? Psychological Bulletin. 131(6), 925-971.

Schkade, D.A. and Kahneman, D. (1998) Does living in California make people happy? A focusing illusion in judgments of life satisfaction. Psychological Science. 9(5), 340-6.
Sen, A. (1992) Inequality re-examined. Oxford: Oxford University Press.

Shields, M. and Wheatley Price, S. (2005) Exploring the economic and social determinants of psychological wellbeing and perceived social support in England. Journal of the Royal Statistical Society. 168(3), 513-537.

Shilton, T., Sparks, M., McQueen, D., Lamarre, M.C. and Jackson, S. (2011) How should we define health? Proposal for a new definition of health. British Medical Journal. 343(7821), d5359.

Stiglitz, J., Sen, A. and Fitoussi, J. (2009) Report by the Commission on the Measurement of Economic Performance and Social Progress. Paris: Commission on the Measurement of Economic Performance and Social Progress.

Taillefer, M-C., Dupuis, G., Roberge, M-A. and Lemay, S. (2003) Health-related quality of life models: Systematic review of the literature. Social Indicators Research. 64(2), 293-323.

Verbeek, M. (2000) A guide to modern econometrics. New York: Wiley.

World Health Organisation. (1948) Constitution of the World Health Organisation. Available at http://apps.who.int/gb/bd/PDF/bd47/EN/constitution-en.pdf [Accessed 26 September 2011].