

Measuring Wellbeing Series

These discussion papers draw from a range of experts on approaches to measuring wellbeing. The intention is to promote debate and thought. Views expressed do not necessarily represent the What Works Centre for Wellbeing.

Richard Layard
December 2016

measuring wellbeing and cost-effectiveness analysis using subjective wellbeing

DISCUSSION PAPER 1

LSE

 what works
wellbeing



About the Measuring Wellbeing Series

The *measuring wellbeing* series consists of **discussion papers** and **how-to guides**. This publication is a discussion paper.

Why ignite discussion about measuring wellbeing?

In our role as an independent collaborating centre and thought leader, the What Works Centre for Wellbeing brings together the disparate theoretical threads to draw out what this means, practically, for decision makers. The Centre recognises wellbeing as a multi-dimensional concept, where a range of definitions and measures may apply and are useful for different purposes.

We don't have just one measure of health or illness; we have many different tools designed to help us understand each in different situations. We'd like to encourage discussion of how different approaches to understanding and measuring wellbeing might be applied as 'the best tool for the job' in different situations as well.

This series of discussion papers includes inputs from leaders in the field. It draws together views of how we could define and measure wellbeing and use this in decision-making in different sectors across UK.

What does this mean for me?

These discussion papers are mainly aimed at analysts, wanting to understand the latest thinking and theoretical underpinnings. However, the accompanying blog and 'Practical Guides' are aimed at all audiences who may be considering how to put wellbeing into practice.

We hope that this paper and series will prompt discussion and bring these methods to life, so we can put these into practice.

Join the discussion about this paper in our online forum

About the What Works Centre for Wellbeing

We are an independent organisation set up to produce robust, relevant and accessible evidence on wellbeing. We work with individuals, communities, businesses and government, to enable them to use this evidence make decisions and take action to improve wellbeing.

The Centre is supported by the ESRC and partners to produce evidence on wellbeing in four areas: work & learning, culture & sport, community, and cross-cutting capabilities in definitions, evaluation, determinants and effects

About Prof Richard Layard

Lord Richard Layard is Emeritus Professor of Economics at the LSE and programme director for Wellbeing at the LSE centre for Economic Performance.

Professor Layard chaired the World Economic Forum Global Agenda Council on Health and Wellbeing, which produced a report at Davos in January 2012, "WellBeing and Global Success". Richard co-edits, with Jeff Sachs and John Helliwell, the World Happiness Reports. He was also a member of the Legatum Commission for Wellbeing Policy chaired by Gus O'Donnell. Its report on Wellbeing and Policy was published in March 2014, and led to the creation of the What Works Centre for Wellbeing.

Professor Layard is the lead investigator of the Centre's Cross-cutting Capabilities team, a consortium of academics and policy makers from LSE, UCL, Oxford, Warwick, OECD, and Action for Happiness.

Acknowledgements

Prof Layard is extremely grateful to Saamah Abdallah and Tessa Peasgood for helpful comments and suggestions.



Contents

1. Common currency proposal	5
2. Exchange rates between currencies	8
3. Wellbeing inequalities	11
4. Cost-effectiveness analysis	14
5. Conclusions	15
Annex 1	16
References	17

Overview

A central problem for decision-makers is how to allocate resources between different activities. When the outcomes of these activities are measured using different units, comparisons are extremely difficult – for example how does one compare a 5% reduction in the rate of burglaries, with a 1% decrease in unemployment rate?

Being able to convert these impacts into the same units allows outcomes to be compared. For this purpose, we propose to use “life-satisfaction” as the common currency. (already there. include:)

This paper proposes a set of exchange rates for converting other measures of wellbeing into equivalent levels of life-satisfaction (section 2); discusses how we could give greater weight to the reduction of misery (section 3); and discusses how policies could be assessed by the wellbeing benefits they bring relative to the net expenditure they involve (section 4).

For all these purposes, what is always needed is information about magnitudes (preferably with confidence intervals). Knowing that something has a significant influence on wellbeing is of little help, unless we also know the size of the influence.



1. Common currency

Criteria

Our main aim is to help decision-makers in government organisations, the third sector, health services, schools and elsewhere make decisions informed by wellbeing evidence. An overarching variable (single or composite) that measures their impact on people's wellbeing will allow them to make coherent choices between competing policies or services. The best variable would have the following characteristics.¹

- It should be salient and meaningful for decision-makers.
- For each member of the population it should comprise a meaningful summary of their quality of life.
- It should have validity and its causes should have been widely studied.
- It should be something that policy-makers can affect.

Proposal

Like the authors of the O'Donnell Report, The World Happiness Report, and the OECD² we believe that life-satisfaction comes nearer to satisfying these characteristics than any other measure (single or composite). The question used to assess life satisfaction in the Measuring National Wellbeing Framework developed by the Office for National Statistics (ONS) is as follows:

“Overall, how satisfied are you with your life nowadays?”

Respondents answer on a scale from zero to ten, where zero means 'not at all satisfied' and ten means 'completely satisfied'.

1. The views of decision-makers

To date, policy-makers have attempted to compare different outcomes by assessing them in units of money. But this only works as a measure of benefit when the preferences of the population can be inferred from private choices (by 'revealed preference').

However, government and charitable activity exist mainly because of public goods and externalities, where private choice reveals little. In these areas policy-makers are in limbo, and wellbeing offers them the right way out. Hypothetical choices (contingent valuations through willingness to pay) are sometimes used, but these have been repeatedly shown by Kahneman and others to produce absurd results.³ What policy-makers need is a direct measure of the impact on wellbeing.

¹ In addition it should be interpersonally comparable, and at least cardinal (or ideally a ratio scale if it is to be combined with length of life – see below).

² O'Donnell et al. (2014) p23; Helliwell et al. (2012), Chapter 2; OECD (2013). See also Brazier and Tsuchiya (2015).

³ Kahneman (2011).



Given the multiple factors that determine people's wellbeing, academics and statistical offices have argued that self-reported assessments are the best way to make an overall assessment of wellbeing. The World Happiness Report contends that such subjective measures 'are arguably the most democratic of well-being measures, since they reflect not what experts or governments think should define a good life, but instead represent a direct personal judgment'

For a policy-maker in a democratic society, hoping to be re-elected, the most relevant measure for each individual is how that individual evaluates his own life. Such an approach has strong face validity – after all respondents are being directly asked to assess their life 'overall'. It is quick and easy to collect data on such an item, and large data sets exist for the UK and worldwide.

By contrast, the hedonic approach to measuring wellbeing by cataloguing positive and negative emotions can be very time-consuming (methodologies such as the Day Reconstruction Method or Experience Sampling), and is subject to large day-to-day variability. Furthermore, questions referring explicitly to 'happiness' have been criticised by commentators on both the left and right as being 'trivial' and momentary. Policy-makers and statisticians seeking to promote subjective wellbeing measures have steered away from the 'h-word' as a result. Eudaimonic approaches to wellbeing, which stress a wide variety of aspects of the good life, including sense of autonomy, meaning or purpose, are appealing but there is currently no agreed set of questions, or indeed elements, of eudaimonic wellbeing.

Evaluative measures such as the life satisfaction question used by ONS ask people to take a step back and reflect on how their life has been recently. This is likely to include how they've been feeling emotionally as well other aspects of life that are important to them (possibly including a sense of control and purpose, etc). If they are an effective assessment, one would expect they might predict future behaviour. This is supported by the analysis of voting behaviour, which suggests that life-satisfaction is a good predictor of election outcomes⁴ (similarly job satisfaction is the best predictor of whether someone will leave their job). Moreover many policy-makers are used to asking people if they are satisfied with their public services. The rational next, more encompassing, step is to ask if they are satisfied with their lives.

2. The views of individuals

But is this an adequate representation of how well an individual is doing in life? One can only say that it is their own evaluation of precisely that.

It will of course reflect many aspects of their experience, which could be captured by direct measurement. These include:

- their momentary feelings experienced over some period, and

⁴ Ward (2015)



- their sense that what they do in life is worthwhile (Eudaimonia)

as measured for example by the four ONS personal wellbeing questions. A composite measure of these answers might have less measurement error for each individual, which is sometimes helpful. But there would be many problems with using a composite variable as our common currency.

The first is that of weighting the components. The weights could be got from a number of different methods:

- statistical techniques such as regressing life-satisfaction on the other variables and using these weights to produce a composite (such that the weights show the importance of each attribute as a determinant of subjective wellbeing).
- asking people directly to weight the importance of the different elements⁵ (such that the weights show how important individuals think each component is – either because they believe it to affect their wellbeing and/or because they believe it to be important in its own right)
- elicited indirectly by a preference-ordering exercise, across options with varying scores on the different wellbeing questions.⁶ (such that the weights can be inferred from the individuals (or populations) choices where, again, they show either the expected impact on subjective wellbeing and/or something of value in its own right).

We lack research evidence on all three approaches. For the regression-based approach simple models risk inadequately capturing the complex relationships between interconnected aspects of someone's day to day feelings and their assessment of their life overall. For example, feelings of happiness or anxiety may be in part caused by an individual's assessment of their life. For the direct and indirect preference weights approaches we have minimal evidence on whether people are able to fully imagine hypothetically described lives, or whether they hold stable preferences towards different aspects of wellbeing.

If a composite measure of wellbeing could be derived, any weighting system would need to be transparent and widely supported. This additional complexity may result in a less clear meaning for decision-makers than a single variable.

Measures of affect and feelings in the moment (or over the last day) inevitably summarise a smaller slice of an individual's life than a question about satisfaction with life 'these days'. This may be appropriate for the question at hand (if say we wish to know about variability in the individual's experience) but in cases where we want a summary of their overall experience over a longer time period it will introduce additional measurement error.

⁵ O'Donnell and Oswald (2015).

⁶ Benjamin et al. (2014).

⁷ Kahneman (2011).



Those researching measures of affect have found that when individuals report their retrospective feelings they differ from those reported in the moment. For example, the length of time spent in a particular emotional state tends to be given less weight in retrospective judgements. This 'duration neglect' leads to differences between measures that cover different time periods. Our preferred measure could allow individuals to give more weight to some time periods over others, reflecting their own value judgements. In reporting life satisfaction the individual can focus on whatever they choose, and can give very little weight to certain periods of time if they choose to do so.

Finally, time-specific hedonic measures have the interesting property that they are much more difficult to explain statistically than life-satisfaction – perhaps because respondents are partly reporting how happy/anxious they are at the specified time compared to how they usually are.⁸ If so, they underestimate the true variance of happiness (or other emotions) in the population. This relates to the next point.

3. Validity

Life-satisfaction has been much studied. It is well-correlated with plausible causes and it is also shown to have dramatic effects upon, for example, life-expectancy.⁹ It has also been found to correlate with other assessments of wellbeing, including neurological assessments.¹⁰

3. Affectability

Comparing life-satisfaction and hedonic measures, it is well-known that life-satisfaction is more affected by external factors like income and education than hedonic measures are. Indeed across countries, income has no significant effect on positive or negative affect (in multiple regressions).¹¹ By contrast, policy-dependent variables like income and employment all have substantial effects on life-satisfaction.

This does not of itself make it a more relevant measure of what policy-makers should be trying to affect. If we really thought that the stream of feelings is all that matters and that hedonic surveys could capture that stream, one should argue for hedonic measures, even if policy-makers resist. But that would require that hedonics become much easier to measure well and thus to explain (e.g. by real-time biological measures). And even then there remains the philosophical question about whether dreams at night count in hedonic experience. (They are a good bit of most people's experience.) So, for the foreseeable future, we propose sticking to life-satisfaction.

Of course, for most activities, it can be important to measure their effects on other variables as well. This makes good sense. But for an overall evaluation of policy options, we need exchange rates that tell us how they translate into life-satisfaction.

⁸ Most hedonic measures are retrospective, involving well-known biases (Kahneman (2011)). By contrast, life-satisfaction is in some sense 'contemporary' and importantly summarises feelings about past, present and future.

⁹ Steptoe and Wardle (2012).

¹⁰ Davidson (1992).

¹¹ Helliwell et al. (2012), p.64.

2. Exchange rates between different wellbeing measures

From what has been said, it is highly desirable that all surveys and evaluations include the question on life-satisfaction. However, for many interventions, especially for health, this measure is not sufficiently context-specific to get accurate measures of effects – or it is simply not available. The question therefore arises as to how results using other currencies can be converted into units of life-satisfaction.

Fortunately the Economic Evaluation Policy Research Unit in Sheffield have examined the relation between life-satisfaction and a whole variety of other measures.¹² They use a variety of surveys in which multiple measures are used (defined in Annex 1). The results are shown in Table 1.

In each case the dependent variable is life-satisfaction, transformed to a scale of 0-10. The table shows in each row how some other variable affects life-satisfaction in different studies. The first column shows the results using standardised variables and the second shows dLS/dX i.e. the change in life-satisfaction (on a scale 0-10) per unit change in the variable in each row (measured in its own units).

The first block of variables are the other three ONS variables as well as wellbeing measured on a visual analogue scale. As can be seen, the correlations are high, except for the question on anxiety (where the responses are highly skewed). Except for anxiety, the results are similar in all the surveys.

The second block are questions on GHQ. Partially because GHQ asks the individual to compare their current state with their usual state, the correlations are much lower. The third block of variables relate to health. The correlations here are mostly in between those on wellbeing measures and those on GHQ. They are also somewhat more variable across surveys.

Finally we turn to a different set of exchange rates: the impact of satisfaction in each domain of life upon overall satisfaction with life. Table 2 (which is somewhat comparable to Figure 1) shows how satisfaction in each domain affects overall satisfaction with life. The analysis by Powdthavee (2012) is a fixed effects analysis multiple regression, using data from the BHPS 1996-2009. Since the results are so similar for men and women, we show the average of the coefficients for men and women. They show the strong potential effect of satisfaction with social variables, and with work, upon overall life-satisfaction. The effect is only potential since the actual effect of, say, a job is the product of how a job affects job satisfaction and how job satisfaction affects life-satisfaction.¹³ Thus Table 2 should be regarded as another set of exchange rates between different wellbeing measures and overall wellbeing (measured by life-satisfaction).

¹² Mukuria et al. (2014).

¹³ There is some upwards bias in these coefficients since a person who is generally satisfied is likely to be satisfied both with life and with its different domains. There are also issues of reverse causation.


Table 1: Impact of different measures upon life-satisfaction (0-10)

	Range of variable	Health improvement and Patient Outcomes		Multi Instrument Comparison		South Yorkshire Cohort over 65		Understanding Society	
		r	dLS/dX	r	dLS/dX	r	dLS/dX	r	dLS/dX
Wellbeing									
Worthwhile (ONS)	0-10	0.80	0.80	0.77	0.84	0.75	0.75		
Happy (ONS)	0-10	0.84	0.84	0.76	0.75	0.80	0.72		
Anxious (ONS)	0-10	0.60	0.53	0.32	0.30	0.50	0.35		
WB-VAS ¹⁴	0-10	0.82	0.490			0.74	0.07		
GHQ									
GHQ	0-36							-0.49	-0.21
GHQ positive	0-18							-0.40	-0.42
GHQ negative	0-18							-0.48	-0.30
Health									
EQ-5D- 5L	(-0.6) - 1	0.63	5.65	0.39	4.56	0.44	3.79		
SF-6D (SF-12) ¹⁵	0.3 - 1	0.72	11.30	0.51	9.15			0.36	5.86
SF-6D (SF-36)	0.3 - 1			0.48	9.22				
WEMWBS ¹⁶	16-70					0.68	0.13		
SWEMWBS ¹⁷						0.66	0.26	0.50	0.25
ICECAP-0 ¹⁸	0-1					0.63	8.77		
ICECAP-A ¹⁹	0-1			0.65	9.71				
EQ-VAS ¹⁵	0-10	0.70	0.80			0.59	0.60		

¹⁴ Visual analogue scale (Four statistics are overrounded).

¹⁵ Short Form 12 or 36 questionnaire

¹⁶ Warwick Edinburgh Mental Well-Being Scale

¹⁷ Shortened Warwick Edinburgh Mental Well-Being Scale

¹⁸ Investigating Choice Experiments Capability Measure for Older people

¹⁹ Investigating Choice Experiments Capability Measure for Adults

Table 2: Domain satisfactions as predictors of life satisfaction, BHPS 1996-2009

Dependent variable: Life satisfaction	r	$\partial LS/\partial X$
Satisfaction with income of household	0.089	0.110
Satisfaction with flat/house	0.063	0.070
Satisfaction with job	0.041	0.086
Satisfaction with amount of leisure	0.055	0.070
Satisfaction with use of leisure	0.146	0.174
Satisfaction with spouse/partner	0.077	0.171
Satisfaction with social life	0.167	0.194
Satisfaction with health	0.138	0.172
Observations	107,501	107,501
R-squared	0.571	0.740

Robust standard errors in parentheses. Fixed effects included.

$p < 0.01$, ** $p < 0.05$, * $p < 0.1$

An obvious question is how QALYs convert into life-satisfaction. There is no satisfactory answer to this question. Instead one has to go back to the scores on the EQ-5D from which changes in QALYs are evaluated. Life-satisfaction has been regressed on the 5 separate dimensions of EQ-5D by Mukuria and Brazier (2013). They find that the relative weight on the different dimensions are quite different from those currently used in QALYs, which is why direct mapping from existing QALYs to life-satisfaction is unsatisfactory. The further issue of how changes in quality of life trade-off with changes in length of life requires further research.

Wellbeing and mental health

What is one doing when measuring the rate at which another variable converts into life-satisfaction? The best approach is to consider the other variable as capturing some but not all of the factors which alter life-satisfaction. For example, as multiple regression shows, mental health is not the only variable that affects life-satisfaction. So mental health and life-satisfaction are not the same thing. On top of that, standard measures of mental health could not ever act as proxies for life-satisfaction as they are highly skewed, with many people having virtually no signs of mental illness. So it is a mistake to argue that any mental health measure is a proper measure of overall wellbeing. It is just one factor which affects it.

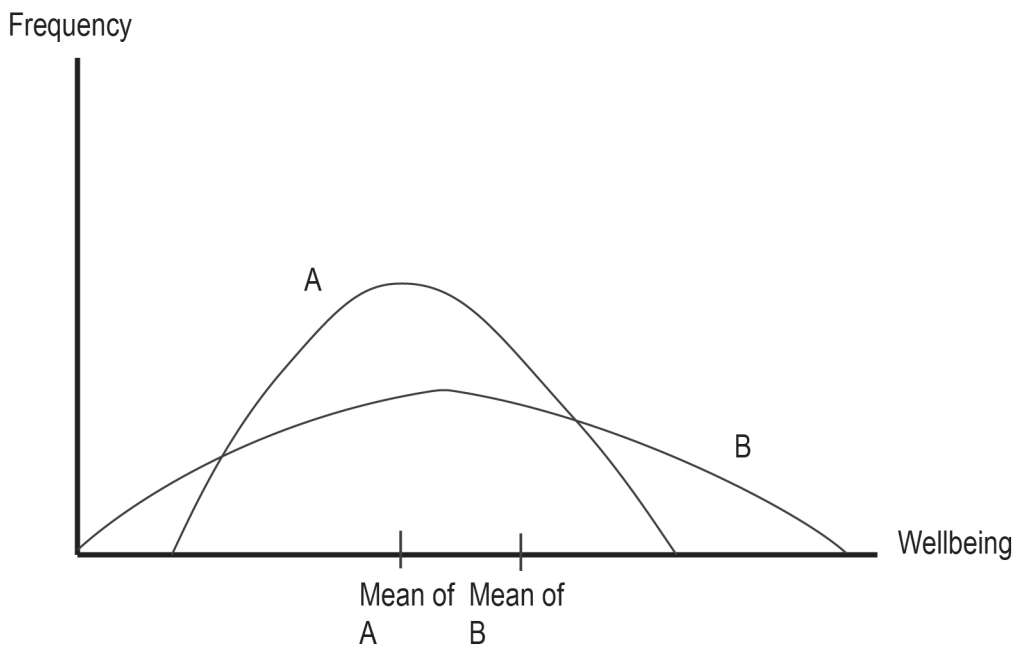


3. Wellbeing inequalities

Once we have measured the wellbeing of every individual, we have a distribution of wellbeing in the population and this distribution embodies a degree of inequality. Figure 1 shows two such distributions (A and B). Two questions then arise

1. The positive question. How do we explain the dispersion of each distribution?
2. The normative question. Which distribution of wellbeing is better – taking into account its unequal distribution, as well as its average?

Figure 1: Two distributions of wellbeing



The positive question

To explain the inequality of life-satisfaction, we need two things. Firstly, we need a system of regression equations which explain why different people have different wellbeing. As an example, the equation depicted in Figure 2 is one such equation. Based on the British Cohort Study, it shows how life-satisfaction at age 34 is affected by aspects of the individual's adult life.

²¹ Layard et al. (2014).

²² Suppose the following regression equation in which all variables are standardised:

Then

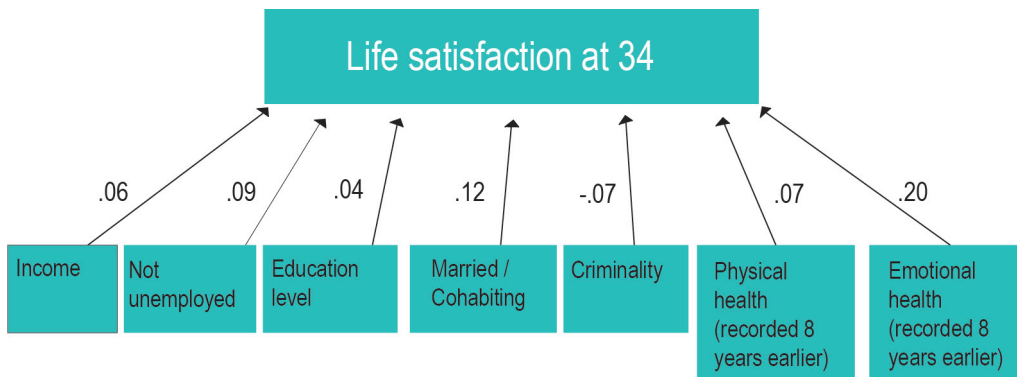
$$W_i = \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + e_i$$

$$R^2 = \beta_1^2 + \beta_2^2 + \dots + \sum_i \sum_j \beta_i \beta_j r_{ij} \quad (i \neq j)$$

Equations like this give us an estimate of how much each measured factor affects the outcome. By looking at which factors matter most, we can see how far equalising each explanatory factor would help towards eliminating wellbeing inequalities.

Figure 2: What are the main immediate influences on adult life-satisfaction?

(Partial correlation coefficients)



So data on cross-sectional wellbeing inequalities can provide important evidence about the causes of wellbeing (including more distant causes). However, whenever possible, such data need to be supplemented by panel data on the same individuals, in order to pin down causality more tightly. And the best evidence of all comes from properly designed experiments.

Such equations help us understand how individual factors such as income and employment status determine wellbeing in a particular context. However, the effects of these individual-level factors may vary from place to place and time to time. For example, individual unemployment status has been found to be a weaker predictor of wellbeing in countries which have higher unemployment rates. Moreover there is a whole range of societal variables which have their own independent effect on the inequality of wellbeing. Little work has been done to explore how these society-level variables predict general wellbeing dispersion.²³ However this brings us directly to the measurement of wellbeing inequalities.

The normative question

If we are worried about the inequality of wellbeing it is because of an ethical concern: we think it would be better if two individuals had the same wellbeing (W) than if one

²⁴ But see Quick (2015) for a review of research on how wellbeing inequality varies between countries. See also Clark et al. (2014) and Stevenson and Wolfers (2008).

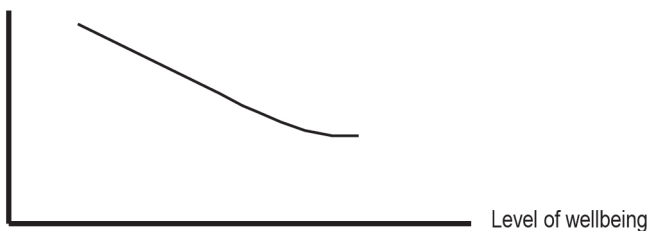


had wellbeing of $(W-\lambda)$ and the other had wellbeing of $(W+\lambda)$. In other words we judge the state of society by an evaluation function $f(W_1, \dots, W_n)$ in which f increases more if a miserable person gains an extra unit of wellbeing, than if a happy person gains an extra unit.

This has implications for policy analysis and for the measurement of wellbeing inequality. For **policy**, it means that when we add up the changes in wellbeing across everyone affected, we need a system of weights, so that extra weight is given to changes in the wellbeing of those who are more miserable. An example is given in Figure 3, where the decision-maker sets out the weights to be applied to changes in wellbeing affecting people at different initial levels of wellbeing. An alternative is to adopt some mathematical formula for determining these weights. For example, one well known formula weights the changes in wellbeing of person i by the weight $1/W_i^\alpha$ where α is positive and reflects the degree to which the decision-maker is averse to inequality.²⁴

Figure 3: Weights to be applied to changes in wellbeing affecting people at different original levels of wellbeing

Marginal value of
extra wellbeing



When it comes to measuring wellbeing inequality, the measure should ideally reflect the same aversion to inequality as is used for policy evaluation. Measures exist which make this possible.²⁵

²⁴ This corresponds to a measure of overall social welfare (Y) given by $Y = \frac{1}{1-\alpha} \sum_i W_i^{1-\alpha}$. In that case $\frac{\partial Y}{\partial W_i} = W_i^{-\alpha}$. This approach is

advocated by Atkinson and Stiglitz (1980) (If $\alpha = 1$, $Y = \sum_i \log W_i$). It requires that wellbeing is measured on a ratio scale.

²⁵ For example, Atkinson (1970) proposed the following approach. Let social welfare be given by the function $Y=f(W_1, \dots, W_n)$. Define W^e as the level of wellbeing which, if everybody had it, would yield the same welfare as $f(W_1, \dots, W_n)$. Then the measure of equality is defined as $E = \frac{W^e}{\bar{W}} \leq 1$. Inequality is $1-E$. This can be made operational using any social welfare function, including the one in the previous footnote.

4. Cost effective analysis

This brings us to the use of wellbeing data for policy-making. For simplicity we can consider decision-makers (in government, health, education or the third sector) as having a given budget to spend. How should they then allocate it, if they wish to maximise overall wellbeing (adjusted for inequality)?²⁶

Evidence

First, they would look at data on the causes of wellbeing (such as that in Figure 2) for ideas on which areas are ripe for further policy development. But such data tells us nothing about the cost-effectiveness of particular interventions. So the second phase is to do controlled experiments; and many have been done already and will be covered in the evidence reviews.

Some interventions are intended to produce immediate results, and the analysis there is relatively straightforward. Others are intended to produce longer-term results. In such cases, the ideal would be to follow-up the gainers and losers for a substantial period of time. For example, if the intervention was in childhood, we would ideally have at least 10 years of follow-up. But this is costly and delays policy-making. So one approach is to use a model of wellbeing over the life-course, which enables one to simulate the long-term net benefit and net costs that can be expected as a result of the immediate changes brought about by the policy intervention.²⁷

In estimating the effects of an intervention, decision-makers need to cover not only the changes in wellbeing of all the parties affected but also the **net financial impact** of the intervention on the budget-holder's budget. For example, a programme of parent training for parents of children with challenging behaviour will have an initial cost which is significantly offset by savings on criminal justice, welfare benefits and special education. Only a model of the life-course makes it possible to estimate such offsetting savings.

The decision criterion

Suppose then that we know the stream of net benefits of a policy (in units of life-satisfaction) and the net costs (in £). How do we know whether to undertake it? It clearly depends on the ratio of net benefits to net costs:

$$\frac{\text{(Present value of net benefits (in LS))}}{\text{(Present value of net budget costs (in £))}}$$

²⁶ See section 3 for possible methods of adjustment.

²⁷ One aim of the Cross-Cutting Team is to estimate such a model (drawing on data for the BCS/NCDS/MCS as well as BHPS/GSOEP/HILDA) which should be available to other strands in due course. Such a model will also help in tracing effects from one generation to another.



This is the bang for the buck. In principle (but not in practice), we should rank all possible policies by this ratio and, starting with the best policy, work down the list until the whole budget is exhausted. In practice a feeling will emerge about what is the critical value which has to be met. For example, NICE have a rough value of the maximum cost per QALY.

In order to calculate the bang for the buck ratio, many key issues have to be addressed – inequalities of wellbeing, the discount rates for life-satisfaction and budget outlays, and the treatment of extra years of life. These were discussed in Wellbeing and Policy (O'Donnell et al, 2014), and the Cross-Cutting Group will be working in conjunction with those responsible for updates to the Treasury Green Book on developing more detailed guidance on all these issues.

The time period of all effects is crucial. This is already embedded in the QALY health methodology, and similar principles apply to any other intervention: we need to translate its effects into a Present Value of changes in life-satisfaction over the whole relevant time period.

It is important to stress that cost-effectiveness can never be the sole guide to decision-making. Many other considerations come into play. But it can be an important factor.

5. Conclusions

The science of wellbeing is in its early stages. But, if wellbeing is the proper objective, we should use all the available knowledge about it. Much of this knowledge has considerable margins of error, but the value of this approach to policy should be judged by comparing it with existing methods, which are generally even weaker. It is better to be roughly right than precisely wrong.

We can end with four main conclusions.

1. Life-satisfaction is the best common currency for policy-makers to use when comparing the outcomes of different interventions. But other measures can be useful, for example to capture finer detail in certain domains.
2. When necessary, other measures can be converted into life-satisfaction.
3. We give more weight to raising life-satisfaction when it is low than when it is high.
4. Policies should be evaluated in terms of the wellbeing improvement (weighted for inequality) per unit of net expenditure from the policy-maker's budget.



Annex 1

The surveys used in Table 1

Health improvement and Patient Outcomes (HIPO): a large UK patient dataset that collected SWB and health data in inpatients recently discharged from hospital in 2014.

Multi Instrument Comparison (MIC): a survey collected from online research panels in 2012 from six countries, including the UK;

South Yorkshire Cohort over 65 (SYC65): a general population sample recruited from a cohort that was recruited from general practitioners in Yorkshire and Humber.

Understanding Society (USoc): wave 1 (2009-10) and wave 4 (2013-14) of the UK household panel.

The variables

Worthwhile (ONS): Overall to what extent do you feel the things you do in your life are worthwhile?

Happy (ONS): Overall how happy are/were you today/yesterday?

Anxious (ONS): Overall how anxious are/were you feeling today/yesterday?

WB-VAS: single item measure of SWB where it focuses on life overall – i.e. “We would like to know how good or bad your life is where “100 means the best life you can imagine and 0 means the worst life you can imagine”. (Table 1 transforms it to 0-10).

EQ-5D-5L: health state classification of five dimensions of mobility, self-care, usual activities, pain/discomfort and anxiety/depression.

WEMWBS: The Warwick Edinburgh Mental Well-being Scale identifies levels of positive mental health in the general population. Asks for time spent in 14 positive states.

SWEMWBS: shortened 7-item version of the WEMWBS. Items on optimism, usefulness, feeling relaxed, thinking clearly, dealing with problems, feeling close to others, and being able to make up one’s own mind.

ICECAP-0 (Investigating Choice Experiments Capability Measure for Older people) and ICECAP-A (Investigating Choice Experiments Capability Measure for Adults): capability measures drawn from Sen’s capability theory. ICECAP-0 has 5 items (attachment, security, role, enjoyment and control). ICECAP-A has 5 items (stability, attachment, autonomy, achievement, enjoyment).

EQ-VAS: visual analogue scale (VAS) which asks respondents “ ... how good or bad their health is on a scale from 0 to worst health to 100 – best health that they can imagine”. (Table 1 transforms it to 0-10).

SF-6D: health state classification system derived from the Short Form 36 or 12. The SF-6D has six dimensions: physical functioning, role limitations, social functioning, pain, mental health and vitality.



References

- Atkinson, A.B. (1970). 'On the measurement of inequality', *Journal of Economic Theory*, 2, pp. 244-263.
- Atkinson, A.B. and Stiglitz, J.E. (1980). *Lectures on Public Economics*, Maidenhead: McGraw-Hill.
- Benjamin, D.J., Heffetz, O., Kimball, M.S. and Szembrot, N. (2014). 'Beyond Happiness and Satisfaction: Toward Well-Being Indices Based on Stated Preference', *American Economic Review*, 104(9), pp. 2698-2735.
- Brazier, J. and Tsuchiya, A. (2015). 'Improving Cross-Sector Comparisons: Going Beyond the Health-Related QALY', *Applied Health Economics and Health Policy*, pp. 1-9.
- Clark, A.E., Flèche, S. and Senik, C. (2014). 'Economic Growth Evens-Out Happiness: Evidence from Six Surveys', CEP Discussion Paper No 1306
- Davidson, R.J. (1992). 'Emotion and affective style: hemispheric substrates', *Psychological Science*, 3(1), pp. 39-43.
- Helliwell, J.F., Layard, R. and Sachs, J., Eds. (2012). *World Happiness Report*. New York: The Earth Institute, Columbia University.
- Kahneman, D. (2011). *Thinking, fast and slow*, London: Allen Lane.
- Layard, R., Clark, A.E., Cornaglia, F., Powdthavee, N. and Vernoit, J. (2014). 'What predicts a successful life? A life-course model of well-being', *The Economic Journal*, 124(F720-738).
- Mukuria, C. and Brazier, J. (2013). 'Valuing the EQ-5D and the SF-6D health states using subjective well-being: A secondary analysis of patient data', *Social Science & Medicine*, 77, pp. 97-105.
- Mukuria, C., Rowen, D., Peasgood, T. and Brazier, J. (2016). *An empirical comparison of well-being measures used in UK. Research Interim Report RR0027: University of Sheffield and the University of York.*
- O'Donnell, G., Deaton, A., Durand, M., Halpern, D. and Layard, R. (2014). *Wellbeing and policy*. London: Legatum Institute.
- O'Donnell, G. and Oswald, A.J. (2015). 'National well-being policy and a weighted approach to human feelings', *Ecological Economics*, 120, pp. 59-70.
- OECD. (2013). *OECD Guidelines on measuring subjective well-being*. Paris: OECD Publishing.
- Powdthavee, N. (2012). 'Jobless, Friendless and Broke: What Happens to Different Areas of Life Before and After Unemployment?', *Economica*, 79(315), pp. 557-575.
- Quick, A. (2015). *Inequalities in wellbeing: Challenges and opportunities for research and policy* New Economics Foundation.
- Stephens, A. and Wardle, J. (2012). 'Enjoying life and living longer', *Archives of Internal Medicine*, 172(3), pp. 273-275.
- Stevenson, B. and Wolfers, J. (2008). 'Happiness Inequality in the United States', *Journal of Legal Studies*, 37, pp. S33-S79.
- Ward, G. (2015). *Is Happiness a Predictor of Election Results?* CEP Discussion Paper No 1343. London, LSE Centre for Economic Performance.

