



To WELLBY or not to WELLBY?

Measuring non-health, non-pecuniary benefits using subjective wellbeing

Joel McGuire, Samuel Dupret, and Michael Plant

This essay was written for the [Worldview Investigations](#) category of Open Philanthropy's [Cause Exploration Prizes](#)

Summary

Open Philanthropy recognises the need to measure benefits beyond health and income. We think that subjective wellbeing is the best tool for the task. Subjective wellbeing (SWB) is measured by asking people to rate how they think or feel about their lives. We propose that the *wellbeing-adjusted life year (WELLBY)*, the SWB equivalent of the DALY or QALY, as the obvious unit to do cost-effectiveness analyses of non-health, non-pecuniary benefits. As our previous work has shown, using WELLBYs can change funding priorities by giving more weight to improving mental health, compared to DALYs or income measures; they may reveal different priorities in other areas.

The advantages of SWB over alternatives are fourfold. (1) SWB captures and integrates the overall benefit to the individual from all of the instrumental goods provided by an intervention. This avoids the challenging problem of assigning moral weights to different goods, makes spillover effects easier to estimate, and clarifies the importance of philosophy. (2) SWB is based on self-reports by the affected individuals whereas Q/DALYs rely on flawed predictions about how good or bad we think an intervention will be for ourselves or others. (3) Using SWB will reveal previously under-captured benefits, such as it has already done for psychotherapy. (4) Measures of subjective wellbeing already exist, are easy to collect, and widely (and increasingly) used in academia and policymaking across a wide range of circumstances and populations of interest. Furthermore, subjective wellbeing measures are reliable and valid instruments and the existing evidence supports consistent use across people.

Having said that, SWB is not without its disadvantages. (1) There is little research on the comparability between SWB scales across people. (2) We don't know where the 'neutral point' lies on SWB scales. (3) We're unsure how to choose the best measure of SWB (e.g., life satisfaction or happiness) or how to convert between them. (4) There are very few cost-effectiveness analyses using WELLBYs. Fortunately, we think these issues can be resolved and we are actively working towards doing so.



1. The problem and a solution

Open Philanthropy's mission is to help others as much as possible. Its human-focused [Global Health and Wellbeing grantmaking](#) aims to save lives, improve health, and increase incomes. However, Open Philanthropy recognises that measuring changes to health or income does not capture all the benefits experienced by the recipients. So, [they ask](#), how should they account for the effects of injustice, discrimination, empowerment, and freedom? To that list, we could also add crime, loneliness, and corruption. Whilst the standardised health metrics, QALYs and DALYs, make it easier to compare different health states in the same units, the broader challenge is to find a common currency that allows sensible trade-offs between health, wealth, and non-health, non-wealth outcomes. How could this be done?

We take it that Open Philanthropy is interested in funding interventions that improve wellbeing. Therefore, reducing discrimination (or injustice, etc.) is good mostly because it increases wellbeing, any other reason is secondary.¹ But what is 'wellbeing'? Philosophers have [three main theories](#): (1) positive experiences, (2) satisfied desires, and (3) a multi-item 'objective list' that includes 'objective' goods such as knowledge, achievement, and love. Conspicuously absent from this list are wealth and health. Most people conclude that, on reflection, these are not *intrinsically* valuable for us (i.e., they are not valuable in themselves). Rather, they are *instrumentally* valuable, a means to achieving some further end. We don't seek money purely for its own sake, but because we think it will make us happier, satisfied, or realise one of the objective goods that plausibly constitutes wellbeing.

What does this mean for Open Philanthropy? Straightforwardly, it suggests we should measure wellbeing directly, if that's possible, rather than any item, or combination of items, that we assume contribute to wellbeing. Can this be done? We think the answer is yes and, indeed, has been lying under our noses for some time, waiting to be put to work.

It is already common practice to combine changes to quality and quantity of health using QALYs and DALYs. The limitations of these measures are well-established (see [Foster, 2020](#)) and there has been a long-standing call for something better and broader than Q/DALYs. What we really want are WELLBYs, *wellbeing*-adjusted life-years.

WELLBYs are constructed by measuring people's *subjective* wellbeing, how they rate the quality of their own lives. One commonly used question, which has been asked in thousands of surveys ([Veenhoven, 2020](#)), is *life satisfaction*: "Overall, how satisfied are you with your life, nowadays?". People are already familiar with the idea of rating our satisfaction in many domains (the jobs we do,

¹ *Welfarism* is the view that wellbeing is the only intrinsic good. *Non-welfarism*, is the view that goods besides wellbeing, perhaps equality and justice, matter intrinsically.



the goods we buy, the services we receive, etc.). We find rating our lives as a whole a small and easy shift².

Once we've selected a subjective wellbeing measure we have to specify how we construct the WELLBYs from it. A straightforward way is to define a WELLBY as a one-point increase in life satisfaction (on a 0 to 10 scale), for one person, for one year. So, if Alice goes from 3/10 to 5/10 for half a year, or from 3/10 to 4/10 for one year, that is worth 1 WELLBY. If Bob is at 7/10, extending their life for one year is worth 7 WELLBYs. We are skirting over some controversial issues here and will return to these in Section 5.

How would the WELLBY capture the impact of health, wealth, discrimination, freedom and the like? Quite straightforwardly, in fact. We can determine, using standard social science research methods, how much each factor impacts subjective wellbeing. So, if a doubling of income and a certain increase in freedom had the same effect on life satisfaction (0.5 WELLBYs for example), we would say they were equally valuable.

Given this, we think the WELLBY is an obvious option, not just for Open Philanthropy, but for anyone else that wants a principled method to compare how different kinds of intervention benefit people and by how much. We do not claim the WELLBY is perfect – it is not – but it does represent a decision-relevant improvement over Q/DALYs and a better alternative to relying on intuitive judgements about the quality of other people's lives.

In the rest of this document we lay out the details of using WELLBYs and SWB. We ask and address the following questions in the rest of this document: How is subjective wellbeing (SWB) measured (Section 2)? How widely has SWB been used (Section 3)? What are the advantages of using SWB compared to the alternatives (Section 4)? What are the challenges of using SWB and what further work is needed to resolve them (Section 5)?

2. How to measure subjective wellbeing

Subjective wellbeing (SWB) is how people rate their feelings or judgements about their lives (i.e., how happy or satisfied they are). SWB is defined by the Organisation for Economic Co-operation and Development ([OECD, 2013](#)) as “good mental states, including all of the various evaluations, positive and negative, that people make of their lives, and the affective reactions of people to their experiences”. SWB is commonly measured by responses to questions like “Overall, how satisfied are

² The median response time for subjective wellbeing questions is less than 30 seconds ([ONS, 2011](#)). And SWB questions have low non-response rates ([Rässler and Riphahn, 2006](#)), and in three of the largest SWB datasets a 10 to 100 times higher response rate than income questions ([OECD, 2013](#)).



you with your life nowadays?” by rating their life on a 0 to 10 scale or “Overall, how happy did you feel yesterday?” (ONS, 2019).

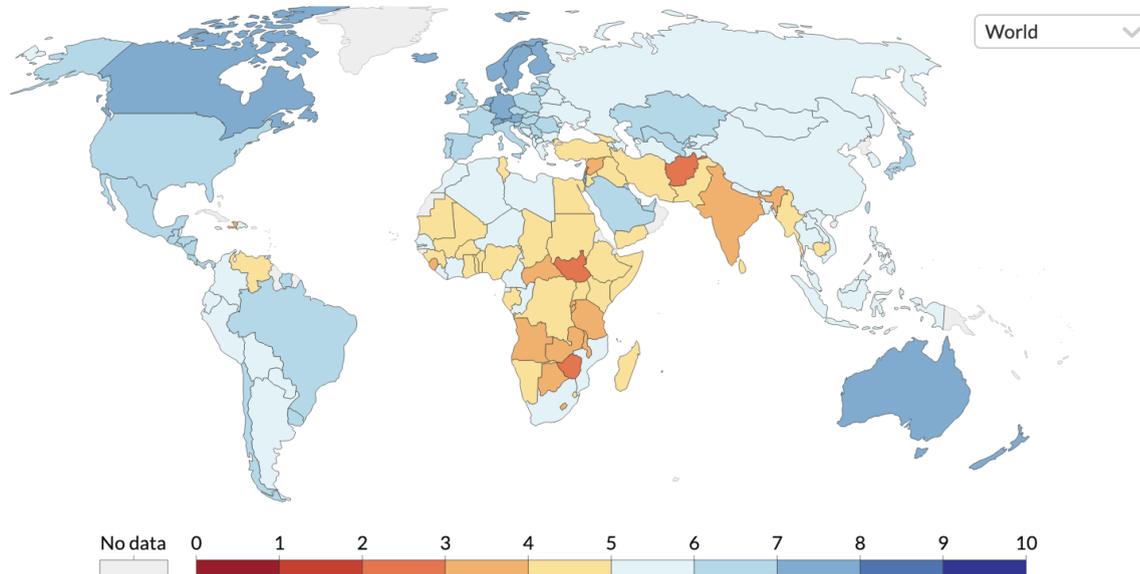
Extensive research has shown that common SWB measures are *valid*. Namely, SWB measures accurately reflect the SWB states that we are trying to measure. To be valid, a measure must be *reliable*; it gives the same consistent output for the same input. Reviews such as the one by the OECD (2013) or Tov et al. (2021) find that common measures of SWB are reliable under these conditions. Beyond reliability, a measure is valid if it captures the underlying phenomenon it set out to capture (i.e., it is correlated with what we think it should be). SWB is, indeed, correlated with the good things in life and negatively correlated to the bad things in life (Kahneman & Krueger, 2006). Relationships, income, and time in nature are positively associated with SWB whilst unemployment, bereavement, commuting, crime, and health problems are negatively associated with SWB (Clark et al., 2018; Dolan et al., 2008). In Figure 1 (based on Gallup World Poll data), we can see that countries afflicted by poverty, lower development, crises, and conflict have lower average life satisfaction compared to richer and more stable nations.

Figure 1. Average life satisfaction across the world in 2020 (Our World in Data)

Self-reported life satisfaction, 2020

“Please imagine a ladder, with steps numbered from 0 at the bottom to 10 at the top. The top of the ladder represents the best possible life for you and the bottom of the ladder represents the worst possible life for you. On which step of the ladder would you say you personally feel you stand at this time?”

Our World
in Data



Source: World Happiness Report (2022)

OurWorldInData.org/happiness-and-life-satisfaction/ • CC BY
Note: The value shown in a given year is the average of that year, the previous year and the following year.

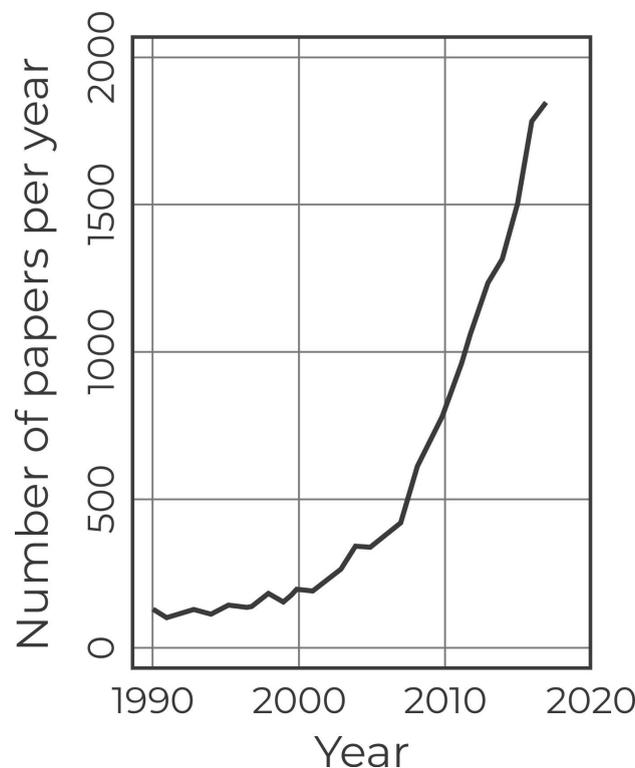


3. Current uses of subjective wellbeing

Subjective wellbeing (SWB) is an accepted method, with official policies, in several countries³ For example, the UK government has guidelines on measurement ([Dolan et al., 2011](#)) and cost-effectiveness analysis ([HM Treasury, 2021](#)). The [Gallup World Poll](#) has surveyed people about their life satisfaction in most countries every year since 2003, the results of which are reported as a cornerstone of the UN's annual [World Happiness Report](#).

The academic field of wellbeing science has grown at a rate of 5.5% per year over the last decades ([Barrington-Leigh, 2022](#)), swelling to probe a panoply of topics. See Figure 2, below, taken from Layard ([2020](#)). Deiner et al. ([2018](#)), coming from a psychology perspective, overviews SWB's relationship to temperament, relationships, performance, creativity and culture. Clark ([2018](#)) reviews the past four decades of happiness economics, covering SWB's relationship to employment, occupation choice, inequality, inflation, and using SWB to value greenery, pollution or noise. Our World in Data ([2017](#)) presents many of the literature's key findings about income, health, culture, and measurement issues.

Figure 2. Number of subjective wellbeing papers over time (from [Layard, 2020](#))



³ Countries with official policies for SWB measurement include Austria, Belgium, Ecuador, Finland, Italy, Israel, Slovenia, and the United Kingdom ([Durand, 2018](#)).



Beyond these typical topics for SWB, it's also been used to estimate the impact of many other events and conditions such as the effect of immigration on movers, natives, and home communities ([Hendriks et al., 2018](#)), the quality of governance ([Helliwell et al., 2018](#)), corruption ([Li & An, 2020](#)), and conflict (Bosnia-Herzegovina: [Shemyakina & Plagnol, 2013](#); Ukraine: [Coupe & Obrizan, 2016](#); Syria: [Cheung et al., 2020](#)).

There is also a literature on the relationship between SWB and hard-to-measure concepts such as discrimination, injustice, freedom and empowerment. For instance, perceived discrimination is strongly related to SWB ($r = 0.24$, based on 328 independent estimates and 144,246 participants, [Schmitt et al., 2014](#)). Affective mental health - which we consider to be a proxy for SWB⁴ - has been used to measure the causal impact of police killings in the USA ([Bor et al., 2018](#)), which are widely considered a consequence of injustice. Differences in freedom across countries better explain differences in SWB than differences in income ([Helliwell et al., 2020](#)), increases in freedom relate to increases in SWB across countries ([Inglehart et al., 2008](#)), and decreases in civil liberties decrease life satisfaction ([Windsteiger et al., 2022](#)). Finally, Fielding and Lepine ([2017](#)) finds that a sense of empowerment has a larger effect on SWB than income.

Wellbeing science is now an established [field](#) and we are not the first to propose WELLBYs. We are not reinventing the wheel here. [Frijters et al., 2020](#), [Birkjaer et al., 2020](#), [Layard & Oparina, 2021](#), and [De Neve et al., 2020](#) have all argued for and demonstrated the usefulness of the WELLBY. However, using WELLBYs to prioritise between policy or philanthropic interventions remains largely unexplored.

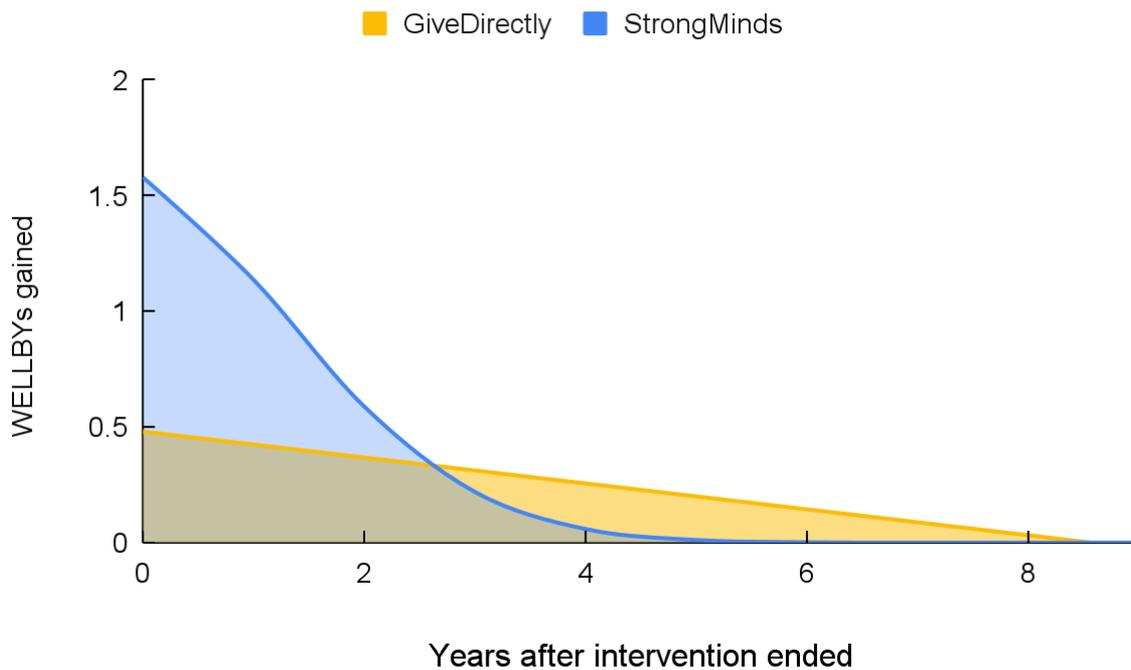
The initial work to synthesise the implications of SWB for policy is promising at this early stage ([Krekel & Frijters, 2021](#); [Global Happiness Policy Report, 2018](#); [Global Happiness Policy Report, 2019](#)), but existing research lacks thorough cost-effectiveness analyses. The Happier Lives Institute was set up to fill this gap and our research has shown how different interventions can be compared using SWB which is applicable in an extremely wide range of circumstances and populations. For example, our research pipeline includes reports on lead exposure, pain relief, immigration reform, digital psychotherapy apps, deworming, and malaria prevention.

⁴ Affective mental health, usually measured with depression scales, involves questions about how people feel, which will directly relate to SWB.



Notably, we have compared the cost-effectiveness of psychotherapy and cash transfers - including long-term effects and household spillovers - with SWB ([McGuire et al., 2022a](#))⁵. We find that [GiveDirectly](#) (a charity that provides cash transfers) produces 7.5 WELLBYs per \$1,000 and [StrongMinds](#) (a charity that provides task-shifted group psychotherapy) produces 71.3 WELLBYs per \$1,000. In Figure 3, we illustrate how the total effect for an individual is the cumulation of the WELLBYs gained over the years.

Figure 3. Total effect for an individual of GiveDirectly and StrongMinds over time



⁵ Our results are currently presented in terms of standard-deviation years of wellbeing gained instead of WELLBYs. This is because we combine data from multiple sources and different SWB measures in a meta-analysis. The typical output for a meta-analysis is standard deviations. One way to convert these results into WELLBYs is to find the typical standard deviation of 0-10 life satisfaction scales. In the [World Happiness Reports](#) the standard deviation is often ~2 points on the 0-10 life satisfaction scale. Therefore, we can convert each standard-deviation year into 2 life satisfaction points per year, namely, 2 WELLBYs.



4. Four advantages of using subjective wellbeing

There are several advantages of using subjective wellbeing (SWB), which we list below.

4.1 SWB captures and integrates the overall benefit to the individual from all of the instrumental goods provided by an intervention (e.g., health, income, empowerment, etc)

If wellbeing is ultimately what matters (i.e., different outcomes are good because they make a person's life good), then an increase in income or empowerment will matter inasmuch as it improves people's wellbeing.⁶ Hence, any benefit from an intervention, be it from health or freedom or a mix of both, should be captured in peoples' reports of their wellbeing. There are three important consequences of this advantage that are worth mentioning.

- (a) With SWB, we don't have to make judgments about the relative moral weights of income, health, empowerment, freedom, or any other good. For more details about how this would be done in practice, see our framework for estimating moral weights using subjective wellbeing ([Donaldson et al., 2020](#)).
- (b) Because SWB is ostensibly measuring what we care about, it's easier to think about modelling and capturing indirect or second-order effects (see [Open Phil's second prompt for the worldview investigation prize](#)). This is why our estimates of household spillovers for cash transfers rely on empirical estimates (see [McGuire et al., 2022a](#)). Similarly, this is a reason why we argue that clearly estimating the long-term benefits of interventions seems important (see [McGuire et al., 2022b](#)).
- (c) Using SWB clarifies the importance of philosophy. We believe that choosing different moral views leads to large changes in cost-effectiveness estimates, as Plant ([2022](#)) recently argued in his philosophical review of Open Philanthropy's cause prioritisation framework. We will expand on this soon by showing that the cost-effectiveness of the [Against Malaria Foundation](#) can change dramatically depending on your philosophical view and the way you operationalize it.

⁶ The alternative for measuring non-income, non-health is to ask people how empowered, free, discriminated, or oppressed they feel or think themselves to be. However, to compare an intervention that increases freedom with one that increases health you would need to compare the value of freedom, health, and other outcomes. But these valuations would be subject to potential bias that SWB avoids.



4.2 SWB is based on self-reports by the affected individuals

Humans make biased predictions about other people's wellbeing or their own future wellbeing ([Coleman, 2022](#)). This can be a problem if evaluators make inferences about how good income and health are for people's wellbeing ([Plant, 2022](#)). This is particularly an issue with existing health measures. The weights assigned to the badness of non-lethal health states (i.e., disability weights) in the Global Burden of Disease's DALYs are derived from judgments made by the general population about health vignettes of diseases ([Global Burden of Disease, 2022](#)). This exposes DALYs to bias stemming from affective forecasting problems (see also a similar issue with QALYs; [Dolan & Metcalf, 2012](#)) and a reliance on healthy people's views as to which states are more healthy than others.

4.3 SWB reveals previously under-captured benefits

Mental health problems like depression are bad because it feels extremely unpleasant to be depressed. Income measures would not capture this and health measures also seem to fail at capturing because they rely on the public's (biased) impression of how bad depression is ([Dolan & Meltcalfe, 2012](#)). Mental health has a large impact on wellbeing, much larger than income, physical health, or unemployment ([Clark et al., 2018](#)). Not only does mental health have a large impact on SWB, mental health treatments can be cost-effective. For example, [StrongMinds](#), a charity which provides task-shifted group psychotherapy in Uganda and Zambia has been rated as cost-effective by Founders Pledge ([Halstead et al., 2019](#)) and we found it to be 9 times more cost-effective than [GiveDirectly](#), a charity that provides direct cash transfers ([McGuire et al., 2022a](#)). Similarly, we expect SWB and WELLBYs to reveal other important causes where the same issues apply. For example, pain, loneliness, and the states of freedom, empowerment, injustice, and discrimination will likely be under-captured by income or DALYs because of the aforementioned reasons.

4.4 SWB is easy to measure and widely applicable

As we mentioned in Section 2, SWB measures already exist, are well studied, and have been applied to a range of topics, circumstances, and populations. Most people in most circumstances can report their subjective wellbeing on a scale without requiring much thought (see footnote 2). Furthermore, measuring SWB outcomes only requires a minimum of a single question that can take less than a minute to answer. That is easier than a consumption survey for a subsistence farmer.⁷ This point may seem trivial compared to other considerations, but the cost-value of

⁷ See the OECD guidelines ([2002](#)) for measuring the food element of consumption alone: "To measure subsistence food production, all items used should be weighed and their origin established at the time meals are being prepared. Since consumption patterns usually vary from one region to another and from season to season, a nation-wide sample of households should be used, with interviews spaced evenly over a full twelve-month period. Surveys of this sort require a fairly large team of trained enumerators and supervisors, and the transport, data processing, and other administrative costs involved may also be considerable."



information matters. Just as we should pursue cost-effective interventions, we should also seek to do research in a way that's maximally informative for the minimum amount of time and resources (see [Lieder et al., 2022](#), for a recent discussion of this).

5. Four challenges for subjective wellbeing and how to solve them

We've argued that subjective wellbeing (SWB) is the best option for measuring non-health, non-pecuniary benefits. It might not be perfect, as we discuss below, but to paraphrase Clark et al. ([2018, chapter 1](#)) we think it's better to have a noisy measure of what really matters than a precise measure of something that matters less.

5.1 Comparability between SWB scales

One potential concern is whether we can compare SWB responses from different people i.e., does Alice's 4/10 mean the same thing as Bob's 4/10?

There is not a lot of research on this topic⁸, but from the work that has been done, we expect people's responses to be comparable. Plant ([2021](#)) argues that people answer scales in a cooperative manner by trying to use scales in the same way as others would. YouGov ([2018](#)) data suggests that people tend to assign the same 0 to 10 scores to different words describing varying levels of 'good' or 'bad'. Work by Kaiser and Vendrik ([2022; Figure 2](#)) found that people use SWB scales approximately linearly, which suggests that people use scales in a similar way. However, some forthcoming work by Benjamin et al. ([2021](#)) finds that people use very different subjective scales *in general* (e.g., how curved is this line on a 0 to 10 scale?) which suggests subjective wellbeing scales will inherit this problem too.

However, even if interpersonal comparability of SWB scales is not perfect there are still a few solutions: (1) We can design SWB scales to better afford interpersonal comparability. For example, we could implement Ng's ([2022, chapter 6](#)) proposal to improve comparability by explicitly including the neutral point as a universal reference. (2) If SWB scale use deviates from comparability in systematic ways, we can mathematically adjust for this.

⁸ The Happier Lives Institute is currently working with Caspar Kaiser and Conrad Samuelsson on a survey to test complex questions about SWB scale use such as linearity, end points, comparability, and the neutral point.



5.2 We don't know where the 'neutral point' lies on SWB scales

The neutral point is where one is neither satisfied or unsatisfied (or neither happy or unhappy). This is equivalent to a DALY of 0, but for wellbeing, this is a point where there is no wellbeing (i.e., one would have the same wellbeing as if they were dead or did not exist). Below this point are states worse than death in wellbeing terms. Knowing this point is crucial for estimating the value of saving a life. This complex area of the WELLBY framework is still being explored and we plan to publish further research on this area later.

5.3 Different measures of SWB might indicate different priorities

Different measures of SWB reflect different philosophies of wellbeing, which means we're uncertain which measure of SWB to prefer. To address this uncertainty, we need to assess how much our priorities change if we use happiness versus life satisfaction⁹ and ensure that SWB data corresponds with all plausible measures of wellbeing. This may mean the creation of new SWB instruments. A related issue is how to convert other SWB measures to our choice metric, but we think this can be solved as a prediction problem by collecting enough data.

5.4 There are almost no cost-effectiveness analyses (CEAs) using WELLBYs

This means that it is hard to tell how WELLBY results differ from previous approaches (e.g., GiveWell's CEAs). Our goal at the Happier Lives Institute is to produce more CEAs using WELLBYs and cultivate a new sub-discipline of wellbeing science to sustain the practice.

6. Conclusion

Subjective wellbeing is a strong contender for measuring non-health and non-pecuniary benefits because it's easy to measure, it captures all perceived benefits, it avoids others telling you how good your life is, there is already an existing (and growing) literature, and it is a reliable and valid way to measure wellbeing. WELLBYs (wellbeing-adjusted life years) provide a coherent framework for cost-effectiveness analysis to assess the value of a wide set of states like freedom, injustice, empowerment, discrimination, poverty, wealth, and health (both mental and physical).

⁹ Boarini et al., (2012) find that shared determinants differ as to the degree of their influence on life satisfaction and affect but mostly have the same sign. In McGuire & Plant (2021) we found that the total effect of transfers differs by 2-13% if we use affective mental health rather than life satisfaction or happiness measures.