We're delighted that HLI's work and the subjective wellbeing approach featured so prominently on the latest 80k podcast with Elie Hassenfeld, the CEO of GiveWell. It was a really high-quality conversation, and kudos to host Rob Wiblin for doing such an excellent job putting forward our point of view. Quite a few people have asked us what we thought, so I've written up some comments.

I split those into four main comments and a number of minor ones. To preview, the main comments are:

1. We're delighted, but surprised, to hear GiveWell are now so positive about the SWB approach; we're curious to know what changed their mind.
2. HLI and GiveWell disagree on what does the most good based on differences of how to interpret the evidence; we'd be open to an 'adversarial collaboration' to see if we can iron out those differences.
3. We'd love to do more research, but we're currently funding constrained. If you – GiveWell or anyone else – want to see it, please consider supporting us!
4. Finally, Rob, we've love to come on the podcast!

However, before we get to *that*, there were some really lovely quotes about our work, and I'd like to share those: [Elie Hassenfeld] "I think the pro of subjective wellbeing measures is that it's one more angle to use to look at the effectiveness of a programme. It seems to me it's an important one, and I would like us to take it into consideration."

[Elie] "...I think one of the things that HLI has done effectively is just ensure that this [using WELLBYs and how to make tradeoffs between saving and improving lives] is on people's minds. I mean, without a doubt their work has caused us to engage with it more than we otherwise might have. [...] it's clearly an important area that we want to learn more about, and I think could eventually be more supportive of in the future."

[Elie] "Yeah, they went extremely deep on our deworming cost-effectiveness analysis and pointed out an issue that we had glossed over, where the effect of the deworming treatment degrades over time. [...] we were really grateful for that critique, and I thought it catalysed us to launch this Change Our Mind Contest. "

Main points

1. We're delighted, but surprised, to hear GiveWell are now so positive about the SWB approach; we're curious to know what changed their mind.

Elie Hassenfeld says the differences in opinion between HLI and GiveWell aren't because HLI cares about SWB and GiveWell does not, but down to differences of opinion interpreting the data[11]. This is great news – we're glad to see major decision-makers like GiveWell taking happiness seriously – but it is also *news* to us!

Listeners of the podcast may not know this, but I (as a PhD student) and then HLI have been publicly advocating for SWB since about 2017 (e.g., 1, 2). I/we have also done this privately, with a number of organisations, including GiveWell, who I spoke to about once a year. Whilst lots of people were sympathetic, I could not seem to interest the various GiveWell staff I talked to. That's why I was surprised when earlier this year, GiveWell made its first public written comment on SWB and was tentatively in favour; Elie's podcast this week seemed more positive.

So, we're curious to know how thinking in GiveWell changed on this. This is of interest to us but I'm sure others would like to know how change happens inside large organisations.

2. HLI and GiveWell disagree on what does the most good based on differences of how to interpret the evidence; we'd be open to an 'adversarial collaboration' to see if we can iron out those differences.

Elie explained that the reason GiveWell doesn't recommend StrongMinds[12] which HLI does recommend, is due to differences in the interpretation of the empirical data. Effectively, what GiveWell did was look at our numbers, then apply some subjective adjustments on factors they thought were off. We previously wrote a long response to GiveWell's assessment and don't want to get stuck into all those weeds here. Elie says – and we agree! – that reasonable people can really disagree on how to interpret the evidence. That's why we'd be interested in an 'adversarial collaboration' to see if we can resolve them. I can see three areas of disagreement.

First, on the general theoretical issue of whether and how to make subjective adjustments to evidence. GiveWell are prepared to make adjustments, even if they're not sure exactly how big they should be. For example, Elie says he's unsure about the 20% reduction for 'experimenter demand effect'. Our current view is to be very reluctant to make adjustments without clear evidence of what size is justified. Our reluctance is motivated by cases such as this: these 'GiveWellian haircuts' can really add up and change the results, so the conclusion ends up being more on the researcher's interpretation of the evidence than the evidence itself. But we're not sure how to think about it either!

A second, potentially more tractable issue, is clarifying what evidence would change our mind about the specific issues. For instance, if there was a well conducted RCT that directly estimated the household spillover effects of psychotherapy in a setting similar to StrongMinds, we'd likely largely adopt that estimate.

A third issue is deworming. Elie and Rob discuss HLI's reassessment of GiveWell's deworming numbers, for which GiveWell very generously awarded us a prize. However, GiveWell haven't commented – on the podcast or elsewhere – on our follow-up work, which finds the available evidence suggests that there are no statistically significant long-term effects of deworming on SWB. This uses the exact same studies that GiveWell relies on; it indicates that GiveWell aren't as bought into about SWB as Elie sounds or they haven't integrated this evidence yet.

3. We'd love to do more research, but we're currently funding constrained. If you – GiveWell or anyone else – want to see it, please consider supporting us!

As a small organisation – we're just 4 researchers – we were pleased to see our ideas and work is influencing the wider discussion about how to do have the biggest impact. As the podcast highlights, HLI brings a different (reasonable) perspective, we have provided important checks on other's work, we provide unique expertise in philosophy and wellbeing measurement, and we have managed to push difficult issues up to the top of the agenda[13][14][15]. We think we 'punch above our weight'.

Elie mentions a number of areas where he'd love to see more research including in SWB and the difficult question of how to put numbers on the value of saving a life.

We think we'd be very well placed to do this work, for the reasons given above; we're not sure anyone else will do it, either (we understand GiveWell don't have immediate plans, for instance). However, we don't have the capacity to do more and we can't expand due to funding constraints. We've only got funding up to October. We need to raise $205k (£162k) to get us to the end of the year, and a minimum of $1,020k (£804k) to fund the next 12 months; ideally, we'd raise more and expand. So, we'd love donors to step forward and support us!

Of course we're biased, but we believe we're a very high leverage, cost-effective funding opportunity for donors who want to see top-quality research that changes the paradigm on global wellbeing and how to do the most good. Please donate here or get in touch at Michael@happierlivesinstitute.org. We're currently finalising our research agenda for 2023-4 (available on request).

4. Finally, Rob, we've love to come on the podcast!

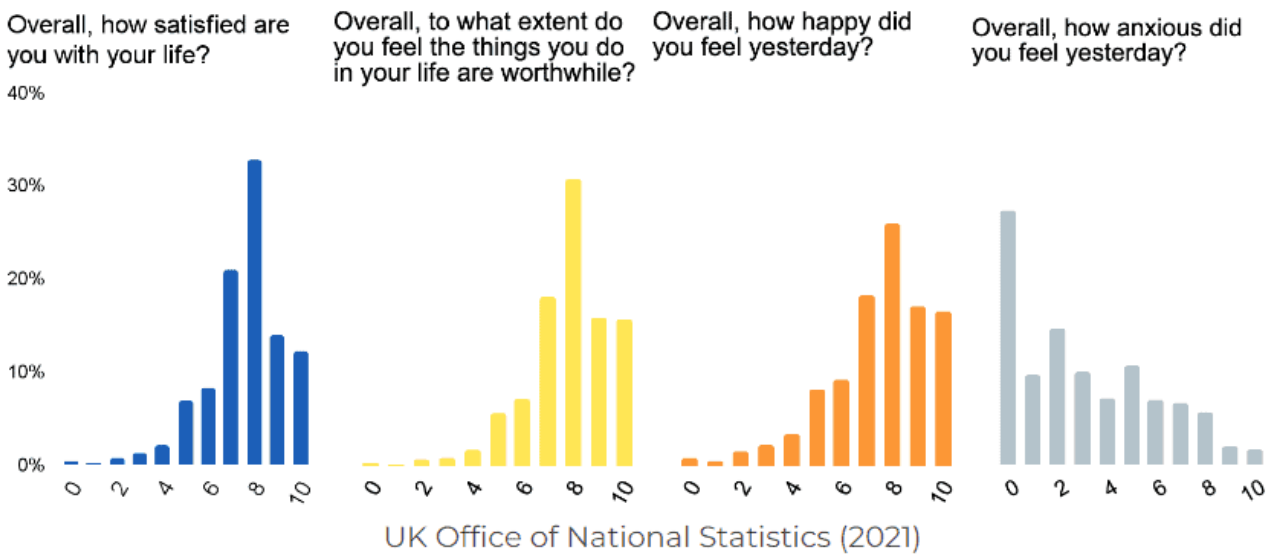
We've got much more to say about topics covered, plus other issues besides: longtermism, moral uncertainty, etc. (Rob has said we're on the list, but it might take a while because of the whole AI thing that's been blowing up; which seems fair).

Minor points

These respond to bits of the discussion in the order they happened.

1. On the meaning of SWB

Rob and Elie jump into discussing SWB without really defining it. Subjective wellbeing is an umbrella term that refers to self-assessments of life. It's often broken down into various elements, each of which can be measured separately. There are (1) experiential measures, how you feel during your life – this is closest to 'happiness' in the ordinary use of the word; (2) evaluative measure are an assess of life a a whole; the most common is life satisfaction; (3) 'eudaimonic' measures of how meaningful life is. These can give quite similar answers. As an example, see the image below for what the UK's Office of National Statistics asks (which it does to about 300,000 people each year!) and the answers people give.



2. On the wellbeing life-year, aka the WELLBY

The way the discussion is framed, you might think HLI invented the WELLBY, or we're the only people using it. That gives us too much credit: we didn't and we're not! Research into subjective wellbeing – happiness – using surveys has been going on for decades now and we're making use of that. The idea of the WELLBY isn't particularly radical – it's a natural evolution of QALYs and DALYs – although the first use seems of the term seems have only been in 2015 (1, 2). The UK government has had WELLBYs as a part of their official toolkit for policy appraisal since 2021.

It is true that HLI is one of the first (if not the first) organisations to actually try to do WELLBY cost-effectiveness; although the UK government has this 'on the books', our understanding is it's not being implemented yet.

3. Is SWB 'squishy' and hard to measure?

Elie: *I think the downside [of measuring SWB], or the reasons not to, might be that on one level, I think it can just be harder to measure. A death is very straightforward: we know what has happened. And the measures of subjective wellbeing are squishier in ways that it makes it harder to really know what it is*

As noted, there are different components of SWB: happiness is not the same thing as life satisfaction. I don't think either of these are *that* squishy or that we don't know what they are; they are different things. I don't think measuring them isn't hard: you

can just ask “how happy are you, 0-10” or “how satisfied are you with your life nowadays, 0-10”! People find it easier to answer questions about their SWB than their income, if you look at non-response rates (OECD 2013). Of course, they are a measure of something subjective, but that’s the whole point. I don’t know how happy you feel: you need to tell me!

4. Does anyone’s view of the good not include SWB?

Elie: *I think some people might say, “I really value reducing suffering and therefore I choose subjective wellbeing.” I think other people might say, “I think these measures are telling me something that is not part of my ‘view of the good,’ and I don’t want to support that.”*

What constitutes someone’s wellbeing, that is, ultimately makes their life go well for them? In philosophy, there are three standard answers (see our explainer here). What matters is (1) feeling good – happiness, (2) having your desires met – life satisfaction, roughly – or (3) objective goods, such as knowledge, beauty, justice, etc, plus maybe (1) and/or (2). It would be a pretty wild view of wellbeing where people’s subjective experience of life didn’t matter at all, or in any way! It might not be all that matters, but that’s another thing.

5. On organisational comparative advantage

Elie: *Because we’re not trying to add value by being particularly good philosophically. That’s not part of GiveWell’s comparative advantage.*

If I can be forgiven for tooting our horn, I do see HLI’s comparative advantage as being particularly philosophically rigorous, as well as really understanding wellbeing science (I’m a philosopher; the other researchers are social scientists). We’re certainly much less experienced than GiveWell at understanding how well organisations implement their programmes.

6. On moral weights

Elie: *I think this is an area — moral weights — where I don't feel the same way. I don't think this is a mature part of GiveWell. Instead, this is a part of GiveWell that has a huge amount of room for improvement*

We could be talked into helping with this! In 2020 we explored how SWB would change GiveWell's moral weights (GiveWell didn't respond at the time). We subsequently have been doing more work on how to compare life-saving to life-improving interventions, including a survey looking at the 'neutral point' and other issues about interpreting SWB data.

7. On the challenges of using SWB given data limitations

Rob Wilbin: *So just the number of studies that you can draw on if you're strictly only going to consider subjective wellbeing is much lower.*

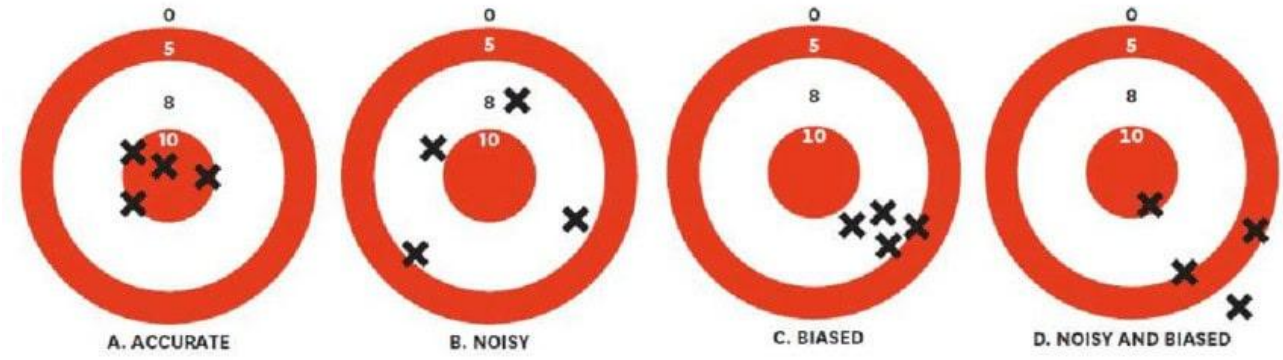
I think another thing that really bites is that subjective wellbeing outcomes are really at the end of the chain of all of these different factors about someone's life — their income, their health, their education, their relationships, all of these different factors.

We see data limitations as the biggest single challenge for SWB. The sort of data we'd want is scarce in low-income countries. We've started to talk to organisation working in LIC and to encourage them to collect data. Our capacity to do this is, however, very limited, but we would expand it if we had the resources. We found enough to compare mental health to cash transfers etc. (and even then we had to convert from mental health scores to SWB scores) but we expect to find much less data to look at other interventions.

On the complex causal chain, part of the virtue of SWB is that, if you have the SWB data, you don't need to guess at how much all the different changes an intervention makes to someone's life affects their wellbeing: you can just look at their answers, and they tell you! Take an education programme. It might change someone's life in all sorts of ways. But if you've done an RCT and measured SWB, you can see the impact without needing to identify where it came from.

Elie: *Maybe there's reason to give credence to the measures that are easier to deal with and easier to know that you've done something good and made someone's life better.*

All this raises a good question: what should we do if we don't have the data we want? As Daniel Kahneman reminds us, we should distinguish two concerns about measurement (and judgement). One is noise, the other is bias.



From Kahneman et al. (2016).

One way to put Elie and Rob's concern is that SWB is a noisy measure. Now, if you have loads of data, you should use a noisy measure over a biased one because all the noise will average out. However, if you don't have much data, and you have a choice between (C on the figure) a non-noisy, biased measure or (B on the figure) a noisy, non-biased one, you could sensibly conclude you'd get closer to the bull's eye with (B).

Here's a hypothetical example that brings this thought. Imagine we're evaluating a new intervention. There is a n = 20,000 RCT that shows it doubles income for pennies. If we convert the income effects to WELLBYs it's much more cost-effective than StrongMinds. But there's an n = 100 RCT with SWB that shows it has a slightly negative but very very imprecisely measured effect on SWB. In this case, I think we'd mostly go with the income evidence (more technically: we'd combine the uncertainty of the income estimate with the conversion to SWB estimate and then combine both the income-converted and the SWB evidence in a bayesian manner).

But the issue is this. How do you know how biased your measures are? You need to establish bias by reference to a 'ground truth' of accuracy – how far are they from the

bull's eye? I'd argue that, when it comes to measuring impact, SWB data is the least-bad ground truth: you learn how important income, unemployment, health etc are by seeing their relationship to SWB. Hence, in the above example, I'd be inclined to go with the income data because there's so much evidence already income does improve SWB! If the full sweep of available data showed income had no effect, I wouldn't conclude that the income data from the hypothetical example was evidence the programme was effective. Of course there will be gaps in our evidence, and sometimes we have to guess, but we should try to avoid doing that.

8. On the tricky issue of the value of saving a life

It's too long to quote in full, but I don't think Rob or Elie quite captured what HLI's view is on these issues, so let me try here. The main points are these; see our report here.

(1) comparing quality to quantity of life is difficult and complicated; there isn't just 'one simple way' to do it. There are a couple of key issues which are discussed in the philosophical literature which haven't, for whatever reason, made it into the discussions by GiveWell, effective altruists, etc.

(2) One of these issues, the one Elie and Rob focus on, is the 'neutral point': where on a 0-10 life satisfaction should count as equivalent to non-existence? We think it's not obvious and merits further research. So far, there's been basically no work on this, which is why we've been looking into it.

(3) how you answer these philosophical questions and make quite a big difference to the relative priorities of life-saving vs life-improving interventions. We got into that in a big report we did at the end of 2022, where we found that going from one extreme of 'reasonable' opinion to the other changes the cost-effectiveness of AMF by about 10 times.

(4) HLI doesn't have a 'house view' on these issues and, if possible, we'll avoid taking one! We think that's for donors to decide.

(5) GiveWell does take a 'house view' on how to make this comparison. We've pointed out that GiveWell's view is at the 'most favourable to saving lives, least

favourable to improving lives' end of the spectrum, and that (on our estimates) treating depression does good more if you hold even slightly less favourable assumptions. This shouldn't really need saying, but philosophy matters!