

Scheffler, Samuel. *Why Worry About Future Generations?*

Oxford: Oxford University Press, 2018. Pp. 160. £16.99 (cloth).

We occupy an unusual point in history. Throughout most of human history, conditions have been relatively stable from one generation to the next. Each generation mostly did the best it could for itself, bequeathing to the next generation an Earth in roughly the same state, save perhaps for a little increase in technological knowledge.

Today, however, it is increasingly recognised that with advancing technology, things we do today can far more dramatically affect both the prospects for the well-being of future generations, and even the chances that those future generations will exist at all (e.g. Derek Parfit, *On What Matters: Volume Two* [Oxford: Oxford University Press, 2011], §127; Olle Häggström, *Here Be Dragons: Science, Technology and the Future of Humanity* [Oxford: Oxford University Press, 2016]; Toby Ord, unpublished). The most-discussed example is climate change: rising temperatures due to industrial-era carbon emissions are likely lead to (for example) increases in extreme weather events, the spread of tropical diseases to currently temperate regions, and significant crop failure, making it much harder for future generations to flourish. Extreme warming, which cannot be ruled out on the basis of current climate science, could even render the Earth simply too hot for human habitation (IPCC, *Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* [Geneva, 2014]).

Climate change, however, is far from the only example. In particular, significantly greater threats of premature human extinction arguably come from biotechnology, and from unaligned artificial intelligence (Nick Bostrom and Milan M. Ćirković, *Global Catastrophic Risks* [Oxford: Oxford University Press, 2008]).

A plausible candidate for the worst *natural* pandemic in history was the Black Death, which is estimated to have killed something like 10% of the world's population in the years from 1347 to 1353. How many people are killed in a given disease outbreak depends on the disease's *virulence* (the proportion of infected persons who die from the infection) and on its *transmissibility* (the strength of the tendency of the infection to pass from one person to another). In the current age of genetic engineering, it is possible, and increasingly easy, to engineer disease strains with increased virulence and transmissibility. If it becomes too easy and regulation does not keep up, it need only take one misanthropic madman to wipe out not a mere 10%, but the entire human race. Large-scale terrorism might no longer require great resources or feats of organisation, and nor need "large-scale" continue to mean fatalities in the thousands (Ali Nouri and Christopher F. Chyba, "Biotechnology and Biosecurity," in Nick Bostrom and Milan M. Ćirković, *Global Catastrophic Risks* [Oxford: Oxford University Press, 2008]).

Artificial intelligence (AI) is potentially a tool for enormous good. Deployed in the right ways, advanced AI could, for instance, lift billions out of poverty. However, as AI becomes more advanced, it also becomes increasingly difficult for humans to supervise and control. We have to ensure that artificially intelligent systems are built with the values we want encoded deep into them, on pain of their enormous power being put to uses quite other than those that anyone intended or could plausibly want. In the extreme, again, unaligned artificial intelligence could even exert such extreme undesirable influence over the course of the world as to render humanity extinct (Nick Bostrom, *Superintelligence: Paths, Dangers, Strategies* [Oxford: Oxford University Press, 2014]).

What should we make of these dizzying and sometimes frankly terrifying prospects?

The natural reaction of any right-thinking person, I think, is that with power comes responsibility. Current people have extremely strong moral reasons to attend to the possible impacts of their actions and inactions on the existence and welfare of future persons. There may be little we could

do (even if we wanted to) to halt the general advance of technology, but we must do everything in our power to ensure that the technology is for good, rather than for ill.

We can spell this natural reaction out more by appealing (at least for definiteness) to a utilitarian axiology. Provided we don't go prematurely extinct, there are, in expectation, *truly enormous* numbers of future people. (A fairly conservative estimate is 10^{16} .) Further, as an evaluative matter, the welfare of all people matters equally ("from the point of view of the universe"). Anything that has any significant adverse impact on average well-being across this enormous future, or anything that has any significant adverse impact on the chance that this potentially glorious future is realised at all, is therefore itself a calamity of correspondingly astronomical proportions. A change that decreases the well-being of all future persons by 0.00001%, or that increases the chance of imminent extinction by 0.00001%, is roughly welfare-equivalent to the intrinsic badness of an event that wipes out 10% of the population throughout the next century (Nick Bostrom, "Existential Risk Prevention as Global Priority," *Global Policy* 4 [2013]: 15-31; Nicholas Beckstead, "On The Overwhelming Importance of Shaping the Far Future" [PhD diss., Rutgers University, 2013]; Hilary Greaves and William MacAskill, "The Case for Longtermism," unpublished).

In this rich and thought-provoking short book, Scheffler asks the titular question: "Why worry about future generations?" He asks it as a philosopher's question, agreeing that we do have powerful reasons to worry, but urging (quite correctly) that it is important to get clarity on precisely what the reasons are.

In particular, Scheffler finds (what I have called) the standard account wanting, largely on motivational grounds. This broadly utilitarian account, he says, misleadingly suggests that the reasons to care about future generations are entirely reasons of *morality* and (further) of moral *duty* (25-6). Further, on that account, moral motivation is generally based on *sympathy* (35). In contrast, Scheffler suggests, there are at least four types of reasons to care about future generations that do not have these features. These are *reasons of interest*, *reasons of love*, *reasons of valuation* and

reasons of reciprocity. These four categories of additional reasons, he suggests, will together constitute an alternative, distinctively non-utilitarian approach (25) to the question.

Reasons of interest arise from the fact that the “value of many of our activities depends... on the survival of humanity long after we ourselves are gone” (44). This is related to the more general tendency that many people have to want to “be part of something larger than themselves” (48). While some are quite content with hedonistic pursuits, and with creating simply for the sake of the creative enterprise, many of us are not like that. We derive far greater satisfaction from engaging in projects that we know to be important to others besides ourselves, and that are collaborative. Thus it is important to the average philosopher, for instance, that she is not merely “a solipsistic scribbler labouring at a private diary” (49). Analogously, Scheffler suggests, it is important to many current persons that there are activities they can engage in that will continue to benefit people after they themselves have died, whether that is a matter of scientific or artistic creation, or social or political action.

Reasons of love arise from the fact that many of us occupy a mental state that deserves to be called a “love of humanity” (62). Just as when we love a person, he suggests, someone who loves humanity is apt to be distressed at the prospect of the object of love being harmed or destroyed.

Reasons of valuation arise because many of the other things that we value, besides humanity itself, would be destroyed if humanity imminently went extinct. For instance, “beautiful singing or graceful dancing or intimate friendship or warm family celebrations or hilarious jokes or gestures of kindness or displays of solidarity” (70).

Finally, Scheffler suggests that despite the fact that physical causation only goes forward in time, we nonetheless stand in an appropriate two-way relation of “genuine mutual dependence” with future generations for considerations of *reciprocity* to arise, since (as above) the value of our own lives depends on the existence of future generations. Scheffler takes the resulting two-way relationship to be relevantly similar to you and I making a habit of each scratching the other’s back.

Scheffler groups the first two of these types of reason together (in chapter 2), and the last two together (in chapter 3). As a matter of conceptual organisation, this grouping seems curious. On my reading, “reasons of reciprocity” is an add-on to “reasons of interest”, while “reasons of love” is a special case of “reasons of valuation”. Still, there is not much in taxonomy; let us turn to matters of substance. How does Scheffler’s four-part account relate to the standard story?

One contrast is that while the standard story is straightforwardly normative, Scheffler’s account is in one clear sense descriptive. He is not pointing out (as the standard account does) that the existence and fate of future generations matters “from the point of view of the universe”, and drawing the conclusion that they *should* therefore also matter to us. Rather, he is rather pointing out that they *do* already matter to us.

This descriptive nature has some practical advantages. If the goal is to get people to care about future generations and act accordingly, a descriptive account that explains to people how they do already so care, when previously they might have been confused on that matter of their own psychology, is likely to have some success. This is an important practical point. At the same time, however, we might be left wondering whether the question has really been answered, or simply shifted. This concern applies, in particular, to Scheffler’s “reasons of love” and “reasons of valuation”. There is not much of a gap between saying that we care about future generations on the one hand, and saying that we are so constituted as to be vulnerable to distress if future generations are harmed or destroyed, on the other. In place of “Why worry about future generations?”, Scheffler’s discussion might shift us to asking instead “why love humanity?” or “why value the continuation of beautiful singing after we are gone?”, but those questions are not so very different.

A second possible contrast is that while the reasons involved in the standard story are clearly ones of morality, Scheffler’s reasons are apparently supposed not to be (except perhaps reasons of reciprocity). This issue looms largest in the context of reasons of interest. On one very natural reading, reasons of interest are reasons of prudence, rather than of morality. We have self-

interested (that is, prudential) reasons to want to live valuable lives, and thus the existence of future generations is instrumentally important to us in prudential terms.

On this prudential reading, paying any significant attention to reasons of interest would be somewhat grotesque. There in the balance hangs an astronomical amount of human welfare: the happiness of quadrillions of future persons, and even the question of whether any of them gets a chance to experience the joys, loves, projects that life has offered us. But never mind *them* – I (rightly) feel better about my own life, as I cycle to work, as I hang out the laundry, if I (correctly) believe that future generations will exist and/or that my activities are benefitting them, and therein lies a significant part of my reason to do anything about it.

Scheffler is of course aware of the prudential reading of “reasons of interest”. Perhaps because of its grotesque aspect, he goes to some length to disavow that reading (53 ff.), framing it as a common *misreading* of his previous book “Death and the Afterlife” (Samuel Scheffler, *Death and the Afterlife* [New York: Oxford University Press, 2013]). Instead, he writes, the intended account is that “[w]e have an interest in [the] survival [of future generations] in part because they matter to us; they do not matter to us solely because we have an interest in their survival” (57).

I find this reply a bit obscure. Insofar as I can understand it, however, it removes any prospect for reasons of interest to provide any *independent* grounds for caring about future generations, over and above either those that are involved in reasons of love and valuation, or those that are involved in the standard story. (On the prudential reading, whatever the demerits of the reasons in question, they are at least independent.)

Whether it is reasons of love and valuation, or instead the familiar utilitarian reasons, that reasons of interest reduce to depend on how seriously we take the “to us” in the above quote. If the grounding fact is that future generations matter *to us*, then reasons of interest seem to add nothing (except the prudential) to reasons of love and/or valuation. The story, on this reading, is that we do in fact value future generations, and as part of this, we place especial value on activities of ours that preserve or

benefit future generations. Alternatively, if the grounding fact is simply that future generations *matter*, as in the standard story, the account appears to be that since we recognise future generations matter, we recognise that activities of ours that preserve or benefit future generations are especially valuable activities. But that adds nothing to the standard story.

By way of aside, one other observation regarding Scheffler's "reasons of interest" is perhaps worth making in passing. Scheffler takes it that we are able to live more valuable lives if future generations exist than if they don't – i.e., it is not only that *conditional on future generations existing*, we lead more valuable lives if we respond to this fact accordingly than if we don't. I have implicitly granted this assumption for the sake of argument, but it is not obviously correct; the issues are similar to some that arise in discussion of intertheoretic and interpersonal comparisons of value and well-being (respectively). Relatedly, while the psychology that Scheffler describes is a common one – wanting not only to be part of something larger than oneself *when things independent of oneself will exist anyway and stand to benefit from one's involvement*, but wanting *to be needed* in the first place – it is not obviously a healthy one. Mindfulness would encourage us to appreciate the intrinsic value in the here and now, turning away from anxiety about whether the here and now is also instrumentally valuable for the elsewhere.

Overall, Scheffler's discussion is interesting and helpful, but I fear that its distance from the standard, broadly utilitarian, account has been oversold. Recall what were supposed to be the main two worries with that account: that it made reasons to care exclusively matters of morality and moral duty, and that it adduced only sympathy as a possible source of motivation. On reflection, we can see that these worries are overstated in the first place.

The utilitarian account does involve exclusively moral reasons to care about future generations. It does not follow, though, that the utilitarian account is exclusively about moral *duty*. What it is more fundamentally about is axiology (a notion towards which Scheffler seems to voice some antipathy (chapter 4)): it centrally involves observations about how enormously much *better* the history of the

world will be if future generations exist and flourish. Coupled with the simple fact (itself fairly widely agreed) that we all have *pro tanto* moral reason to promote the good, these observations are enough to establish that we have moral reasons, and potentially very powerful ones, to take steps to protect and benefit future generations. It is an optional further step to conclude from that that we have moral *obligations* to take those steps, and (even if we do take that step) it is an optional tactical decision to choose to emphasise the obligations rather than merely the reasons. Indeed, many utilitarians themselves are not particularly interested in the notion of moral obligation, or (as in the case of scalar utilitarianism) explicitly eschew any such notion.

The reasons offered by Scheffler's own account seem to fall into three categories (I am not sure which reasons fall into which categories, and one or more categories might be empty). First, some of them might be prudential reasons. This would make them genuinely independent of moral considerations, and perhaps motivationally useful in the case of people who don't feel a strong pull towards moral concerns; but Scheffler himself vehemently denies this reading. Second, they might not be *reasons* to care at all: they might simply be observations that we do care, coupled with some more concrete remarks about the nature of our caring (as seems to be the case for "reasons" of love and of valuation). Third, some of them might be moral reasons, although not necessarily matters of obligation. This is perhaps the most charitable reading, but in that case the reasons under discussion are of a piece with those involved in the standard, broadly utilitarian account.

The picture that seems to emerge is: Future generations matter. Because we recognise that and because we are somewhat decent creatures, they also matter *to us*: we love humanity, we value the things whose continuation would make the lives of future people good (as they have made our lives good), and we take particular satisfaction in activities that protect the existence and interests of future people. Reasons of interest, love and valuation are a real part of the story, but, on this reading, they all stem from the fundamental, broadly utilitarian, evaluative facts. They do not, after

all, constitute a distinctively non-utilitarian approach. They need not invite a reading in terms of moral obligation, but neither need the utilitarian account.

Suppose we adopt this reading. What then of motivation? The motivations arising from love, valuing and satisfaction are not, or not directly, matters of sympathy. So, fortunately if the motivation of sympathy is a shaky one, there are additional motivational resources to draw on, beyond those of sympathy. But this just shows that the idea that utilitarian motivation is purely a matter of sympathy is anyway an artefact of a misleading and impoverished moral psychology. Some utilitarians (e.g. J.J.C. Smart, "An Outline of a System of Utilitarian Ethics," in J.J.C. Smart and Bernard Williams, *Utilitarianism: For and Against* [Cambridge: Cambridge University Press, 1973], 7, 31) have perhaps suggested that idea, but most (e.g. John Stuart Mill, *Utilitarianism* [Oxford: Oxford University Press, 1998], chapter 2) have had more sophisticated accounts all along.

Hilary Greaves, Faculty of Philosophy, University of Oxford