# Average Utilitarianism Implies Solipsistic Egoism

**Abstract**

Average utilitarianism and several related axiologies, when paired with the standard expectational theory of decision-making under risk and with reasonable empirical credences, can find their practical prescriptions overwhelmingly determined by the minuscule probability that the agent assigns to solipsism—i.e., to the hypothesis that there is only one welfare subject in the world, viz., herself. This either (i) constitutes a reductio of these axiologies, (ii) suggests that they require bespoke decision theories, or (iii) furnishes a novel argument for ethical egoism.

Average utilitarianism (AU) holds that the overall value of a world is equal to the average lifetime welfare of all welfare subjects in that world.[1] Among population axiologies, AU has some notable and distinctive virtues: For instance, it avoids Parfit's "Repugnant Conclusion" (Parfit, 1984), and would be chosen by selfish agents from behind a particularly natural version of the veil of ignorance. It has plenty of vices to counterbalance these virtues (see Hurka (1982a), among many others), and has therefore never been an especially popular doctrine. But it has attracted its share of advocates over the years, including Hardin (1968), Harsanyi (1977), and Pressman (2015).[2]

In the world we appear to inhabit, AU has the slightly dispiriting consequence that we can only make a very small difference to the overall value of the world: Since the number of welfare subject is very large, even acts that produce enormous welfare improvements in absolute terms have only minuscule effects on *average* welfare.

[1]This, at least, is an especially natural and widely-discussed version of average utilitarianism. There are others, which we will come to shortly.

[2]Mill (1863) can also be read as an average utilitarian (see Gustafsson, forthcoming, fn. 2), though the textual evidence for this reading is not entirely conclusive.

To illustrate: There are currently about $7 \times 10^9$ human beings alive on Earth. There are also many billions of mammals, birds, and fish being raised by humans for meat and other agricultural products. And there are perhaps some $10^{11}$ mammals living in the wild, along with similar or greater numbers of birds, reptiles, and amphibians, and a significantly larger number of fish—conservatively $10^{13}$, and possibly far more.[3] This is despite a significant decline in wild animal populations in recent centuries and millennia as a result of human encroachment.[4]

To determine the *total* number of welfare subjects (by which we can divide total welfare to find an overall average), we must consider past as well as present individuals. (Future individuals count too, of course, but their numbers are much harder to estimate, and may depend on our choices.) Estimates of the number of human beings who have ever lived are on the order of $10^{11}$ (Kaneda and Haub, 2018). But this number is dwarfed by past populations of non-human animals. In wild animal populations, most individuals die young (with smaller animals being both more numerous and shorter-lived), so birth and death rates in the wild animal population as a whole are unlikely to be less than 1 per individual per year (roughly corresponding to an average lifespan of 1 year).

Being extremely conservative, then, we might suppose that all and only mammals are welfare subjects and that $10^{11}$ mammals have been alive on Earth at any given time since the K-Pg boundary event ($\sim$ 66 million years ago), with a population birth/death rate of 1 per individual per year. This implies a "timeless population" of at least $\sim 6.6 \times 10^{18}$ welfare subjects. Being a bit less conservative (though perhaps still objectionably conservative), we might suppose that all and only vertebrates are welfare subjects and that $10^{13}$ vertebrates have been alive on Earth at any time in the last 500 million years (since shortly after the Cambrian explosion), with a

[3]See Tomasik (2019) and citations therein. These numbers come with very large error bars, but are more likely to be underestimates than overestimates, and in any case, the exact numbers will not matter much for the arguments that follow.

[4]For instance, Smil (2013, p. 228) estimates that wild mammalian biomass has declined by 50% in the period 1900–2000 alone.

population birth/death rate of 10 per individual per year. This implies a timeless population of at least $\sim 5 \times 10^{22}$ welfare subjects.

Even by the more conservative estimate, we find that providing one unit of welfare to one individual increases average welfare by at most $\frac{1}{6.6\times10^{18}} \approx 1.5 \times 10^{-19}$ units. By the less conservative estimate, increasing someone's welfare by one unit will increase average welfare by at most $\frac{1}{5\times10^{22}} = 2 \times 10^{-23}$ units.

There is, however, one hypothesis according to which the total number of welfare subjects in the world is quite a bit smaller: solipsism. Solipsism is the proposition that only I exist (or, in your case, the proposition that only you exist), and that what appears to be an external world populated with other individuals is in fact just a figment of my (resp. your) imagination. If solipsism is true, then the total number of welfare subjects, the size of the universal population, is one.

Solipsism is surely improbable. But just how improbable is it? Like average utilitarianism, it has several notable virtues: It is simple and ontologically parsimonious. It is a natural conclusion to draw from various arguments for external world skepticism (of the kind found, e.g., in Descartes (1641) and Berkeley (1710)), if one does not simultaneously accept some "rescue" hypothesis like theism. (And it is at the very least contentious whether those arguments for external world skepticism have ever been satisfactorily refuted.) Solipsism also provides a powerful answer to the otherwise intractable question "Why am I me?"—namely, "There isn't anyone else I could have been!" And finally, it is a recurring and enduring idea in the history of philosophical thought, having been entertained (in various forms) by such thinkers as the Buddhist philosopher Ratnakirti (see Kajiyama (1965)), Wittgenstein (1922), and Hare (2009).[5]

For our purposes, it will be necessary to go beyond these general observations

[5] It is hard to find straightforward endorsements of the strongest form of metaphysical solipsism, but this provides little if any evidence against the view: First, true solipsists have little reason to publicize their beliefs. Second, and more to the point, if solipsism is true and the whole external world is a figment of my imagination, there is no obvious reason to expect that I would imagine the particular sub-figments that I mistake for other people to go around proclaiming that the whole world is a figment of *their* imagination.

and say something about what *probability* one might reasonably assign to solipsism. A bit more specifically, the interestingness of the following arguments depends on the claim that one's credence in solipsism should be *not absurdly small*—not less than, say, $10^{-9}$ (one in a billion). Of course, assigning probabilities to philosophical hypotheses is at best a matter of rough guesswork—we do not, for instance, have objective chances to go by. But we can do better than simply pulling plausible-seeming numbers out of thin air. We can, for instance, consider a more complete and fine-grained space of possibilities over which our probabilities should sum to 1, and aim for reflective equilibrium among the credences we assign to the various possibilities in that space. Fully carrying out that exercise here would be a tedious experience for the reader. But we can do a first approximation, aiming simply to find a plausible lower bound.

It seems clear that my credence that my commonsense view of the world has "got things basically right", metaphysically speaking, should not be greater than 0.9. That is, given how little we have to go on, the lack of expert consensus in basic metaphysics, and trying to correct for the general human tendency toward overconfidence, I should have at least 0.1 credence that the world is in *some* way fundamentally very different than I take it to be. A good chunk of that credence should go to "some possibility that nobody has ever thought of". But it also seems overconfident, conditional on my ordinary view of the world being wrong, to have credence greater than 0.9 in *that* possibility. This leaves at least 1% of my credence to distribute over known revisionary metaphysical hypotheses. And then to get some sense of a lower bound on my credence in solipsism, I should ask first, "Are there any other known revisionary hypotheses that are *many orders of magnitude* more probable than solipsism?" (to which, it seems to me, the answer is "no") and second, "Are there thousands or millions of known revisionary hypotheses that are at least roughly as plausible as solipsism?" (to which the answer again seems to be "no"). Taken together, these observations suggest that my credence in solipsism should be at most a few orders of magnitude less than 0.01.

All in all, then, while it would strike me as somewhat unreasonable to assign solipsism a probability greater than $10^{-2}$, it also seems unreasonable to assign it a probability less than $10^{-9}$. To assign any more extreme probability would not display due modesty about our understanding of matters metaphysical.[6]

Now, consider an average utilitarian, Ava, who assigns solipsism a subjective probability of $10^{-9}$, and must choose between taking one unit of welfare for herself, or providing a thousand other welfare subjects with a thousand welfare units each. And let's suppose she believes that, if solipsism is false and the external world/other minds are real (hereafter, "realism"), then the total number of welfare subjects in the world is $10^{18}$. (For simplicity, I am rounding down our already conservative lower-bound estimate of $6.6 \times 10^{18}$, and ignoring the credence Ava ought to have in larger population sizes, which would only strengthen our conclusions.) And let's assume (without loss of generality) that whether solipsism is true or false, average welfare prior to Ava's intervention is 0. This situation is summarized in Table 1.

Now, suppose that Ava responds to risk in the standard way, by maximizing expected value. Given the facts stipulated above, the expected value of the altruistic option is $(1-10^{-9}) \times 10^{-12} + 0 \times 10^{-9} \approx 10^{-12}$ while the expected value of the selfish option is $(1-10^{-9}) \times 10^{-18} + 1 \times 10^{-9} \approx 10^{-9}$. That is, even though the altruistic option *almost certainly* yields a million times more value than the selfish option, the selfish option has a thousand times greater *expected* value, because if solipsism is true and only Ava exists, then Ava can have astronomically greater impact on

[6]In drafting this paper, I become curious what credence people actually *do* assign to solipsism, when prompted. So, in an *extremely* unscientific and non-incentive-compatible survey on social media ($n = 32$), I asked respondents to assign credences to solipsism defined as "the hypothesis that only I exist (or, in your case, the hypothesis that only you exist), and that what appears to be a physical world containing objects, other people, etc is in fact just a product of/contained within my (or, in your case, your) mind". Among those who gave sharp credences, responses spanned the range $[0, 0.5]$, with a median of $10^{-2}$ and a mean of $\sim 0.105$. Arbitrarily excluding the answers I take to be clearly irrational (those outside the interval $(0, 0.1]$) still gives a median of 0.01, but reduces the mean to $\sim 0.048$. On the other hand, including a few participants who gave interval credences, at the lower bound of their intervals, gives a median of $10^{-6}$ and a mean of $\sim 0.091$. Of course, all of these numbers are quite a bit greater than $10^{-9}$.

| | **Solipsism is false.** | **Solipsism is true.** | **EV** |
|---|---|---|---|
| | $1 - 10^{-9}$ | $10^{-9}$ | |
| Altruistic option | $\frac{10^6}{10^{18}} = 10^{-12}$ | 0 | $\sim 10^{-12}$ |
| Selfish option | $\frac{1}{10^{18}} = 10^{-18}$ | 1 | $\sim 10^{-9}$ |

Table 1: Solipsistic swamping for average utilitarianism

average welfare than she could otherwise hope for. Despite the enormous disparity in stakes, we find that Ava ought to choose the selfish option as long as her credence in solipsism is greater than $\sim 10^{-12}$. Conversely, holding fixed her credence in solipsism at $10^{-9}$, we find that she should give her own interests a billion times more practical weight than anyone else's—i.e., her interests carry a billion times greater weight in expectation.[7]

We assumed that Ava accepts a particular (very natural) version of average utilitarianism, which has been our exclusive focus so far. But as Thomas Hurka has emphasized, there are many non-equivalent views that can be described as "average utilitarian". In Hurka (1982a,b), he describes a total of eleven such views, which he names A1–A11. (Ava's view, which I have called AU and which tells us to maximize average lifetime welfare in the timeless population, is Hurka's A1.) These theories are not all equally vulnerable to solipsistic swamping. For A2, which tells us to maximize the sum of momentary welfare averages (i.e., averaging welfare at each time and then summing across times), the crucial number that determines how much solipsism magnifies one's efficacy is the size of the *present* population, rather than the *timeless* population. For A7, which tells us to maximize the average lifetime welfare of *present and future* people (ignoring the past), the crucial number is of course the size of the present and future population. So either of these views somewhat dampen the swamping phenomenon. On several other views (Hurka's

[7]As I discovered while revising this paper, the vulnerability of AU to this sort of "solipsistic swamping" has been noticed once before, in a short blog post by Caspar Oesterheld (Oesterheld, 2017). I am not aware of any other discussions of the phenomenon, or any discussion of its generalization to the other axiologies besides AU discussed below.

A3, A4, A6, A8, A9, and A11), which involve averaging across times, things depend on how long the agent believes she will exist if solipsism is true, and how long the Universe as a whole will contain welfare subjects if solipsism is false. A5 and A10 evade the solipsistic swamping problem entirely—but, as Hurka points out, these views are independently very implausible.

Solipsistic swamping also threatens other axiologies that try to capture the intuitive attractions of AU in large-population contexts (e.g., to avoid the Repugnant Conclusion). For instance, consider the view that Hurka (1983) calls "Variable Value I" (VV1), according to which the value of a population $\mathbf{X}$ is given by

$$V(\mathbf{X}) = \overline{\mathbf{X}} f(|\mathbf{X}|)$$

where $\overline{\mathbf{X}}$ is the average welfare level in $\mathbf{X}$, $|\mathbf{X}|$ is the number of welfare subjects in $\mathbf{X}$, and $f$ is a function that is strictly increasing, strictly concave, and has a horizontal asymptote.[8] "Variable value" axiologies are meant to resemble total utilitarianism for small populations and average utilitarianism for large populations, reflecting the intuition that adding more (happy) individuals to a population adds value when the population is small, but has diminishing marginal value as population size increases.

How vulnerable is VV1 to solipsistic swamping? Very roughly, the crucial factor is the ratio $r$ between $f(1)$ and the horizontal asymptote of $f$. If this ratio is much larger than the minimum population size conditional on realism, then VV1 may agree arbitrarily closely with total utilitarianism, and so be safe from solipsistic swamping. If $r$ is much smaller than that minimum population size, then VV1 will reduce the extreme practical weight that AU given to solipsism by approximately a factor of $r$. To illustrate the latter case, suppose that

$$f(|\mathbf{X}|) = 1 - \frac{1}{1 + 10^{-9}|\mathbf{X}|}$$

[8]This view is also discussed by Ng (1989), under the name "Theory X".

| | **Solipsism is false.** | **Solipsism is true.** | **EV** |
|---|---|---|---|
| | $1-10^{-9}$ | $10^{-9}$ | |
| Altruistic option | $\frac{10^3}{10^{22}} \times f(10^{22}) \approx 10^{-19}$ | 0 | $\sim 10^{-19}$ |
| Selfish option | $\frac{1}{10^{22}} \times f(10^{22}) \approx 10^{-22}$ | $1 \times f(1) \approx 10^{-9}$ | $\sim 10^{-18}$ |

Table 2: Solipsistic swamping for Variable Value I

Here $r \approx 10^9$, meaning that the "axiological weight" given to larger populations converges to roughly one billion times the weight of a singleton population.

Now the problem of solipsistic swamping persists, but is much less extreme. Since $\frac{f(10^{18})}{f(1)} \approx 10^9$, the relative weight of the solipsistic hypothesis is reduced by a factor of nearly $10^9$, and so using the numbers from our original example (in Table 1), we now find that the altruistic option has greater expected value than the selfish option. But the problem is far from vanquished. Consider a new agent, Valerie, who (i) accepts VV1 with the $f$ specified above, (ii) assigns solipsism a credence of $10^{-9}$, (iii) accepts our slightly-less-conservative estimate of the minimum population size conditional on realism, which for simplicity we will round down to $10^{22}$ (as compared to $10^{18}$ in the case of Ava), and (iv) must choose between taking one welfare unit for herself or providing a thousand other welfare subjects with one welfare unit each (for a thousand units in total, as compared to a million in the case of Ava). For Valerie, like Ava, selfishness is the order of the day—even though she is nearly certain that the altruistic option will produce far more value (Table 2). A bit more generally, given a VV1 axiology with the $f$ specified above, $10^{-9}$ credence in solipsism, and a minimum population of $10^{22}$ conditional on realism, Valerie should give her own interests at least 10,000 times as much weight as anyone else's, because of her credence in solipsism.[9]

[9]Hurka also describes a few he calls Variable Value II, that applies an increasing and concave transformation $g$ to average welfare, so that the overall value of a population $\mathbf{X}$ is given by $g(\overline{\mathbf{X}})f(|\mathbf{X}|)$. This view behaves like VV1 for our purposes, with the additional caveat that, if $g$ is sufficiently concave or the amount by which an agent can improve her own welfare sufficiently great, $g$ will further moderate the difference in stakes between solipsism and realism.

On the other hand, the *rank-discounted utilitarian* (RDU) axiology defended by Asheim and Zuber (2014) faces a more extreme form of solipsistic swamping than even AU. On this view, the value of a population $\mathbf{X}$ is given by

$$\sum_{r=1}^{|\mathbf{X}|} \beta^r w_{\mathbf{x}}(x_r)$$

where the members of $\mathbf{X}$ are indexed in order of increasing welfare by their rank $r$, and $\beta \in (0, 1)$ is a constant that determines the degree to which worse-off individuals are prioritized over better-off individuals.

This view does not uniformly discount the interests of each individual in large-population scenarios, as average utilitarianism does—the worst-off individual, for instance, always gets exactly the same weight regardless of population size. But because the weight given to the interests of better-off individuals diminishes exponentially with their welfare rank, the interests of *all but the very worst off* can be dramatically discounted in large-population scenarios. For instance, suppose that Ragnar (i) accepts RDU with $\beta = .99999$ ($\beta$ closer to 1 implies less discounting of the better off), (ii) assigns solipsism a credence of $10^{-9}$, (iii) believes that there are at least $10^{18}$ welfare subjects, conditional on realism, and (iv) must choose between taking one welfare unit for himself or proving a million other welfare subjects with one welfare unit each. Further, suppose Ragnar knows that none of the million individuals he has the chance to benefit are among the $10^9$ worst-off. (If the total number of welfare subjects is at least $10^{18}$, then it is *extremely* unlikely, in any given choice situation, that one is in a position to help any of the $10^9$ worst-off.) On the other hand, Ragnar recognizes that if solipsism is true, then he is very well positioned to improve the welfare of the very worst off individual in the whole Universe—namely, himself.

Because the weight given to an individual's welfare shrinks exponentially with their rank, Ragnar will find that there is a truly dramatic disparity between the solipsistic and non-solipsistic stakes (Table 3): If solipsism is true, then he can do

| | Solipsism is false. | Solipsism is true. | EV |
|---|---|---|---|
| | $1 - 10^{-9}$ | $10^{-9}$ | |
| Altruistic option | $< 10^6 \times \beta^{10^9}$ | 0 | $< \sim 10^{-4337}$ |
| Selfish option | $< 1 \times \beta^{10^9}$ | $\beta$ | $\sim 10^{-9}$ |

Table 3: Solipsistic swamping for RDU ($\beta = .99999$)

roughly $10^{4343}$ times more good (by acting selfishly) than he could do if solipsism is false (by acting altruistically). RDU, then, produces a much stronger form of solipsistic swamping than even AU.[10]

So, not just AU but several other *prima facie* plausible population axiologies as well are vulnerable (in different degrees) to solipsistic swamping. What are we to make of this? We might say that solipsistic swamping is just another instance of expected value calculations being dominated by tiny probabilities of extreme scenarios (as, for instance, in Pascal's wager (Pascal, 1669), Pascal's mugging (Bostrom, 2009), or the St. Petersburg game (Bernoulli, 1738)). This is true as far as it goes, but serves only to categorize the problem, not to solve it. Perhaps the solution is to follow the perennial suggestion of simply ignoring very small probabilities (Buffon, 1777; Smith, 2014; Monton, 2019), but this approach has quite serious drawbacks and is generally considered unsatisfactory (Hájek, 2014; Isaacs, 2016; Lundgren and Stefánsson, forthcoming).

Moreover, solipsistic swamping is just the limiting case of a more general phenomenon, viz., that when combined with standard expectational decision rules, average utilitarianism, variable value views, and rank-discounted utilitarianism all seem to over-weight small-population scenarios. For instance, consider an average utili-

[10] Of course, the proponent of RDU could always hand-select a $\beta$ close enough to 1 to avoid solipsistic swamping. But apart from the *ad hoc*-ery of such a move, it seems very likely that such a large $\beta$ would make RDU practically indistinguishable from total utilitarianism. At the very least, RDU must thread a very tight needle to avoid collapsing into egoism on the one hand or totalism on the other. (And whatever $\beta$ we choose, as long as we hold it fixed in the face of new empirical information, we will be in constant danger of collapsing into solipsistic egoism if we come to believe that the total world population is significantly larger than we thought, or into de facto total utilitarianism if we come to believe that it is smaller.)

tarian who assigns 1% credence to the hypothesis that the Universe will only ever contain $10^{20}$ welfare subjects, and 99% credence to the more optimistic hypothesis that advanced future civilizations will eventually support $10^{50}$ welfare subjects or more (Bostrom, 2013). The same absolute welfare improvement matters $10^{30}$ times more in the former scenario and therefore, discounting for her credence, matters $10^{28}$ times more in expectation. Thus, even though she is quite confident in the "optimistic" hypothesis, she should premise her choices almost entirely on the "pessimistic" hypothesis.[11] More generally, she will end giving almost no practical weight to states that imply a very large population, even when those states are very probable. Apart from optimism about the future of humanity, such states might correspond to (i) hypotheses that attribute sentience to more beings, e.g., to insects, other invertebrates, or relatively simple artificial intelligences or (ii) cosmological hypotheses that imply that the Universe is very large and hence contains many non-Earth-originating welfare subjects (as well as exobiological hypotheses that imply a higher probability of welfare subjects emerging in a given star system). If we find this general phenomenon of "small-population swamping" counterintuitive, then ignoring small probabilities won't help, since we cannot assume that small-population scenarios will always deserve *de minimis* probabilities.

Maybe the conclusion to draw is that some population axiologies cannot be combined with standard decision theory, but must be equipped with their own, bespoke theories of decision-making under risk that avoid the tyranny of small-population scenarios. It not immediately obvious what these decision theories should look like, and in departing from standard decision theory, they are likely to incur significant theoretical costs.[12] But in any case, if we conclude that certain views

[11]Of course, this is complicated by the facts that (i) if the optimistic hypothesis is true, agents like us may be able to have much greater impact on *total* welfare, and so perhaps a similar level of impact on average welfare, and (ii) we may be in a position to significantly influence the population size of future civilization.

[12]Here is one example: Teruji Thomas suggests an extension of average utilitarianism that ranks risky prospects by *expected total utility divided by expected population size* (Thomas, 2016, p. 150). This view straightforwardly avoids solipsistic swamping. But it has the very significant downside of violating statewise dominance—that is, preferring options that yield worse outcomes in every

in population ethics cannot safely appeal to the best developed and most widely accepted theory of decision-making under risk, this on its own would be a notable conclusion.

Absent some clever decision-theoretic escape, we are left with a conditional: If average utilitarianism, a variable value view, or rank-discounted utilitarianism is correct, then the best thing we can do, *ex ante*, to make the world a better place is to act selfishly (to greater or lesser extents, depending on the axiology). This leaves us, of course, with two further options: Reject all these axiologies, or embrace (*de facto*, impartially motivated) ethical egoism.

# References


Asheim, G. B. and S. Zuber (2014). Escaping the repugnant conclusion: Rank-discounted utilitarianism with variable population. *Theoretical Economics 9*(3), 629–650.

Berkeley, G. (1982 (1710)). *A Treatise Concerning the Principles of Human Knowledge*. Indianapolis: Hackett.

Bernoulli, D. (1954 (1738)). Exposition of a new theory on the measurement of risk. *Econometrica: Journal of the Econometric Society 22*(1), 23–36.

Bostrom, N. (2009). Pascal's mugging. *Analysis 69*(3), 443–445.

Bostrom, N. (2013). Existential risk prevention as global priority. *Global Policy 4*(1), 15–31.

Buffon, G.-L. L. d. (2010 (1777)). Essai d'arithmetique morale (Essays on moral


possible state of the world. (As proof: Consider a choice between a lottery L1 that yields one individual with welfare 10 in state S1 or 9 individuals with welfare 20 in state S2, and lottery L2 that yields 9 individuals with welfare 11 in S1 or 1 individual with welfare 21 in S2, where S1 and S2 are equiprobable. L2 statewise dominates L1, by average utilitarian lights, but the "expected total utility divided by expected population size" method of evaluating lotteries gives L1 a value of 19 and L2 a value of 12.)

arithmetic). In A. Ockenfels and A. Sadrieh (Eds.), *The Selten School of Behavioral Economics: A Collection of Essays in Honor of Reinhard Selten*, pp. 245–282. Heidelberg: Springer. Translated by John D. Hey, Tibor M. Neugebauer, and Carmen M. Pasca.

Descartes, R. (2016 (1641)). *Meditations on First Philosophy*. Cambridge Texts in the History of Philosophy. Cambridge: Cambridge University Press. Edited and translated by John Cottingham.

Gustafsson, J. E. (forthcoming). Our intuitive grasp of the repugnant conclusion. In G. Arrhenius, K. Bykvist, and T. Campbell (Eds.), *The Oxford Handbook of Population Ethics*. Oxford: Oxford University Press.

Hájek, A. (2014). Unexpected expectations. *Mind 123*(490), 533–567.

Hardin, G. (1968). The tragedy of the commons. *Science 162*(3859), 1243–1248.

Hare, C. J. (2009). *On Myself, and Other, Less Important Subjects*. Princeton, NJ: Princeton University Press.

Harsanyi, J. C. (1977). Morality and the theory of rational behavior. *Social Research 44*(4), 623–656.

Hurka, T. (1982a). Average utilitarianisms. *Analysis 42*(2), 65–69.

Hurka, T. (1982b). More average utilitarianisms. *Analysis 42*(3), 115–119.

Hurka, T. (1983). Value and population size. *Ethics 93*(3), 496–507.

Isaacs, Y. (2016). Probabilities cannot be rationally neglected. *Mind 125*(499), 759–762.

Kajiyama, Y. (1965). Buddhist solipsism – a free translation of Ratnakirti's Samtanantaradusana. *Journal of Indian and Buddhist Studies (Indogaku Bukkyogaku Kenkyu) 13*(1), 435–420.

Kaneda, T. and C. Haub (2018). How many people have ever lived on earth? Population Reference Bureau. First published in 1997, updated in 2002, 2011, and 2018. Accessed 22 November 2019. URL: https://www.prb.org/howmanypeoplehaveeverlivedonearth/.

Lundgren, B. and H. O. Stefánsson (forthcoming). Against the de minimis principle. *Risk Analysis*.

Mill, J. S. (1863). *Utilitarianism*. London: Parker, Son, and Bourne.

Monton, B. (2019). How to avoid maximizing expected utility. *Philosophers' Imprint 19*(18), 1–25.

Ng, Y. (1989). What should we do about future generations? Impossibility of Parfit's Theory X. *Economics and Philosophy 5*(2), 235–253.

Oesterheld, C. (2017). The average utilitarian's solipsism wager. Accessed 20 March 2021. URL: `https://casparoesterheld.com/2017/03/15/the-average-utilitarians-solipsism-wager/`.

Parfit, D. (1984). *Reasons and Persons*. Oxford: Oxford University Press.

Pascal, B. (1852 (1669)). *Pensées*. Paris: Dezobry et E. Magdeleine.

Pressman, M. (2015). A defence of average utilitarianism. *Utilitas 27*(4), 389–424.

Smil, V. (2013). *Harvesting the Biosphere: What We Have Taken from Nature*. Cambridge, MA: The MIT Press.

Smith, N. J. J. (2014). Is evaluative compositionality a requirement of rationality? *Mind 123*(490), 457–502.

Thomas, T. (2016). *Topics in Population Ethics*. Ph. D. thesis, St. Cross College, University of Oxford.

Tomasik, B. (2019). How many wild animals are there? First published 2009, updated 7 August 2019. Accessed 15 November 2019. URL: https://reducing-suffering.org/how-many-wild-animals-are-there/.

Wittgenstein, L. (1999 (1922)). *Tractatus Logico-Philosophicus.* Mineola, NY: Dover. Translated by C. K. Ogden.