IMPOSSIBLE ETHICS: DO POPULATION ETHICAL IMPOSSIBILITY RESULTS SUPPORT MORAL SKEPTICISM AND/OR ANTI-REALISM?

BY

VICTOR MOBERGER

Abstract: In this paper, I discuss two different metaethical challenges based on population ethical impossibility results. According to the *anti-realist* challenge, the results pose a serious threat to the existence of objective moral facts. According to the *skeptical* challenge, the results pose a serious threat to the reliability of our moral intuitions. My aim is to systematically explore and evaluate these challenges. In addition to clarifying the issues, I argue that population ethical impossibility results *do not* in fact support any anti-realist or skeptical conclusions.

1. Introduction

The field of *population ethics* studies various ethical issues raised by our relations to future generations. When taking a stand on such issues, we often need to compare the intrinsic value, or social welfare, of possible future populations from an ethical point of view, and in doing so, we need to rely

on certain claims or assumptions about what matters in such comparisons. For example, is equality with respect to the distribution of individual well-being important? How important is it in comparison with other factors, such as the average level of individual well-being, or the number of individuals? According to various population ethical *impossibility results* (or *impossibility theorems*), there are deep tensions between our intuitions about these matters. More specifically, the results purportedly show that certain intuitively plausible, or even compelling, adequacy conditions for ranking populations in terms of social welfare cannot be reconciled.[1]

Over the past four decades or so, there has been much discussion within population ethics about the implications of these results, not least with respect to the prospects of developing a viable and comprehensive population axiology. While this literature is primarily focused on first-order issues, it also contains several interesting but usually brief suggestions concerning *metaethical* implications, that is, implications concerning the semantics, metaphysics, and epistemology of moral thought and discourse in general. For example, Jeff McMahan writes:

Problems in [population ethics] seem to me the most difficult and intractable of all the problems of which I am aware in normative and practical ethics. They suggest that it is a real possibility that any moral theory that is both complete and coherent will have implications that are intuitively intolerable. It is these problems, therefore, rather than arguments in metaethics about the queerness of objective values, the connections between normativity and motivation, and so on, that seem to me to pose the greatest challenge to realism in ethics.[2]

Here, the idea is that the kind of choice that population ethical impossibility results would foist upon us—between incompleteness, incoherence, and intuitive intolerability—poses a greater threat to moral realism than even the combined strength of the standard metaethical arguments for anti-realism.

Similarly, Larry Temkin agrees with Shelly Kagan that inconsistencies brought out by population ethical impossibility results

would be *much* easier to accept if we abandon realism in ethics (or never had it). After all, if morality, etc., is just a tool that we use to serve various purposes, it might be that the best such tool is an inconsistent model! (If we aren't describing an independent realm, why think that the best tools are consistent?)[3]

Again, the idea is that the impossibility results threaten moral realism by being much more easily accommodated by some form of anti-realism

---

[1] The *locus classicus* of population ethics is Parfit (1984, Part 4). Impossibility results are found in, for example, Ng (1989), Blackorby and Donaldson (1991), Carlson (1998), and Arrhenius (2000a, 2000b, 2011, forthcoming).

[2] McMahan (2013, p. 34).

[3] Kagan, personal correspondence with Temkin, quoted in Temkin (2012, p. 521, n. 42), emphasis in original.

(although Temkin reports that neither he nor Kagan is prepared to abandon realism on this ground).[4]

A different metaethical challenge is presented by Gustaf Arrhenius, who comments on one of his own population ethical impossibility theorems as follows:

> The […] theorem shows that […] at least one of our considered moral beliefs is false. Since consistency is, arguably, a necessary condition for moral justification, we would thus seem to be forced to conclude that there is no moral theory which can be justified.[5]

Here the idea is that population ethical impossibility results pose a threat of moral *skepticism* rather than anti-realism. What is directly threatened is moral justification rather than moral reality, as it were. While Arrhenius does not himself push this skeptical worry, he does present it as a significant threat to be overcome.

My aim in this paper is to systematically explore and evaluate these two population (meta)ethical challenges, thus moving beyond the brief remarks found in the existing literature. I will begin with the skeptical challenge and then move on to the anti-realist challenge. In addition to clarifying the issues, I will argue that population ethical impossibility results *do not* in fact support any skeptical or anti-realist metaethical conclusions.

The paper is structured as follows. In Section 2, I unpack the skeptical challenge, filling in several details and bringing out its underlying assumptions. In Section 3, I argue that, once these details are worked out, the challenge loses its apparent force. In Section 4, I move on to the anti-realist challenge. I argue that it too fails, and for similar reasons. I end in Section 5 with a brief summary of the main conclusions.

2. *The skeptical challenge*

To illustrate the skeptical challenge, let us consider the following adequacy condition from one of Arrhenius' impossibility theorems:

The Non-Sadism Condition: An addition of any number of people with positive welfare is at least as good as an addition of any number of people with negative welfare, other things being equal.[6]

---

[4]Temkin (2012, p. 521, n. 42). See also Fleurbaey et al. (2009), pp. 274, 284). A more specific anti-realist challenge is due to Christopher Cowie (2022), who appeals to a 'fixed points' approach to moral concepts in order to argue that population ethical impossibility results reveal conceptual incoherence, which he in turn takes to support a moral error theory. Another more specific anti-realist challenge is due to Gustaf Arrhenius, who suggests that population ethical impossibility results 'might give support to some form of non-cognitivism' (Arrhenius, forthcoming, p. 398). I will return to these specific challenges in Section 4.

[5]Arrhenius (2011, p. 23). See also Arrhenius (2000a, pp. 200–201; 2004, p. 214; forthcoming, pp. 392–393), Temkin (2012, § 14.8), and Cowie (2022, § 2.1).

[6]Arrhenius (forthcoming, p. 95).

At first blush this seems hard to deny. How, after all, could it be *better* to add people who are badly off rather than people who are well-off?

Suppose that a number of seemingly inescapable conditions like this can be shown to be mutually inconsistent. According to the skeptical challenge, this shows that our *best* efforts at approaching moral truth are deficient, which, in turn, suggests that *all* our efforts at approaching moral truth are deficient. Here is Arrhenius:

> The idea behind the sceptical conclusion is that if even moral intuitions formed under ideal circumstances can be showed to be epistemically unreliable, then we have reasons to be sceptical about the evidential value of *all* moral intuitions. Since the moral beliefs underpinning the adequacy conditions in the theorems are as considered and scrutinized as one can ask for and thus can be considered to be formed under ideal circumstances, and since the theorem shows that at least one of them must be false, the antecedent of the sceptic's argument is fulfilled, or so the argument goes. Hence, we cannot contain our scepticism only to moral beliefs in population ethics. Rather, the theorems throw a shadow of doubt over all of our considered moral beliefs.[7]

Thus, the idea is that the moral intuitions in play are epistemically *privileged* in the sense that they are formed under ideal circumstances, and if even such privileged intuitions are unreliable or unjustified, then we should conclude that the same goes for all moral intuitions (or at least enough of them to warrant the label moral skepticism). How, after all, can we trust our moral judgment on any issue, when it has been shown to lead us into such unannounced error? To use an analogy, it is as if we found out that the propositions $1 + 1 = 2$ and $2 + 2 = 4$ are mutually inconsistent, and consequently that at least one of them has to be false. If so, we would presumably begin to distrust our arithmetical intuitions more generally. Similarly, if we found out that *modus ponens* and *modus tollens* are in fact irreconcilable, we would be at our wits' end.[8]

This challenge is simple and yet potentially formidable. Compared with other skeptical challenges in moral epistemology, the population ethical challenge has a significant advantage in being able to sidestep several contested issues. At present, debates over moral skepticism are centered on two major issues, concerning the skeptical import of *moral disagreement* and *debunking explanations*. These debates are by now thoroughly intertwined with highly general, and highly entrenched and convoluted, epistemological issues, such as the nature and epistemic significance of peer disagreement, as well as the seemingly intractable issue of coming up with counterexample-free and yet sufficiently precise and substantive accounts of justification, reliability, and other central epistemic notions. By contrast, if our basic and *shared* moral intuitions can be shown to be straightforwardly *inconsistent*, then the skeptical import of disagreement and debunking could perhaps be set to one side. As

---

[7] Arrhenius (forthcoming, p. 393), emphasis in original.
[8] I borrow the arithmetical analogy from Cowie (2022, p. 282).

the imagined arithmetical and logical cases suggest, inconsistencies between shared and strongly held intuitions could potentially pose a significant skeptical threat, and without mention of their genealogy.

Despite these initial promises, however, it remains to be seen whether the population ethical case is indeed analogous to the arithmetical and logical scenarios in relevant respects. So let us spell things out in a bit more detail.

To begin with, what is the precise aim of the challenge? In the above quote, Arrhenius formulates the challenge as targeting the *reliability* of (all) moral *intuitions*. He also talks of the *evidential value* of moral intuitions, and in adjacent passages he talks instead of moral *justification*. He never mentions moral *knowledge*, however. Also, he appears to use the terms 'moral intuition' and 'moral belief' more or less interchangeably.

Below, I will put things in terms of the reliability of moral intuitions, but nothing important turns on this precise way of speaking. I take it that if all of our moral intuitions should turn out to be epistemically unreliable, then this would apply also to our moral beliefs. I also take it that once we have established the unreliability of our moral intuitions/beliefs, then this would spell doom for moral justification, moral knowledge, and anything else in the vicinity.[9] Thus, the potential skeptical implications are broad, even if formulated more narrowly. (The analogous diagnosis also seems like the right one in the imagined arithmetical and logical cases.)

Also, in keeping with the above-mentioned neutrality with respect to contested epistemological issues, I will take the notion of reliability as intuitive. I think this is innocuous. In our imagined arithmetical and logical meltdown scenarios, we don't need to know precisely what reliability amounts to know that we face a serious skeptical worry, and I take it that the same goes for the envisaged population ethical scenario.

In order to evaluate the population ethical skeptical challenge, it will be helpful to look more closely at one of our imagined skeptical scenarios. I will focus on the arithmetical case. This case looks like a 'good' one, in the sense that it does seem to have skeptical import. Once we have spelled out the workings of that skeptical scenario in more detail, we will know what to look for in the population ethical case.

Here is a more formal reconstruction of the skeptical argument in the arithmetical case:

Arithmetical meltdown
P1. $1 + 1 = 2$ and $2 + 2 = 4$ are mutually inconsistent.
P2. If $1 + 1 = 2$ and $2 + 2 = 4$ are mutually inconsistent, then certain privileged arithmetical intuitions are mutually inconsistent.
P3. If these privileged arithmetical intuitions are mutually inconsistent, then these privileged arithmetical intuitions are unreliable.

[9]Cf. Joyce (2006, pp. 216–217).

P4. If these privileged arithmetical intuitions are unreliable, then arithmetical
    intuitions are generally unreliable.

C. Thus, arithmetical intuitions are generally unreliable.

More informally put, we move from *privileged inconsistency* to *privileged unreliability*, and then from *privileged unreliability* to *general unreliability*.

A crucial issue with respect to clarifying the argument is how the notion of epistemic privilege should be understood. For the argument to work (setting aside the obviously false but imagined-to-be-true P1), the notion of epistemic privilege needs to be understood in such a way that P2, P3 and P4 all come out as (sufficiently) plausible. More specifically, we want to make the notion strong enough to render P3 and P4 plausible, but not so strong that P2 is called into doubt. So, with this constraint in mind, how should we understand the notion?

To begin with, we can note that we are very *confident* that the propositions in question are true. Moreover, this confidence is *tried and tested*. The arithmetical propositions have been scrutinized throughout the ages by the best minds that humanity has to offer, and no one has found any grounds for doubting them. This mirrors what Arrhenius claims for the relevant population ethical intuitions, that they are 'as considered and scrutinized as one can ask for and thus can be considered to be formed under ideal circumstances'.[10]

While this *scrutinized compellingness* plausibly provides a necessary condition for the requisite epistemic privilege, it is not sufficiently strong to render an inconsistency between the relevant intuitions a threat to their reliability. In other words, scrutinized compellingness is not sufficiently strong to render P3 in the above argument plausible.

To see this, note first that among a set of mutually inconsistent but individually compelling propositions, one of them might be rejectable in a way that does not endanger the others. This may be because the considerations or intuitions supporting it are different from the considerations or intuitions supporting the others, so that by rejecting it we avoid contamination effects. Moreover, this rejectability will be enhanced if the proposition in question is also significantly less compelling than the others. And once we have such a rationale for singling out one of the propositions for rejection, we can use the conjunction of the others as premises in a logically valid argument against it, thus further reinforcing its status as rejectable. This move will not be available if the propositions involved are *equally* compelling, where the term 'equally' signifies that there is no rationale for singling out any particular proposition for rejection. And this

----

[10]Arrhenius (forthcoming, p. 393).

is presumably the case with respect to $1 + 1 = 2$ and $2 + 2 = 4$; they are equally compelling in this sense.[11]

Epistemic privilege as *scrutinized equal compellingness* is still not sufficiently strong, however. To see this, consider *lotteries*. In a typical lottery, it is rational to assume of each ticket, considered individually, that it will lose. If we add these assumptions up, however, we get an inconsistency with the known fact that at least one ticket will win. Thus, we have a series of individually highly plausible propositions:

Ticket 1 will lose.
Ticket 2 will lose.
Ticket 3 will lose.
…
Ticket *n - 1* will lose.
Ticket *n* will lose.

However, the very setup of the lottery guarantees that at least one of these propositions—let us call them *losing-claims*—is false. Hence, the setup constitutes an 'impossibility theorem' with respect to the losing-claims: they cannot all be true.

Now, suppose that the lottery contains a trillion tickets, exactly one of which is a winner. If so, each of the losing-claims, considered individually, will be extremely plausible, and so for any given such claim we can be extremely confident that it is true. Moreover, this confidence will survive any scrutiny we can muster, and it applies equally to each losing-claim. Thus, scrutinized equal compellingness is fulfilled. Still, the inconsistency between these claims does not seem to have any skeptical import. Consider the analogous skeptical argument in the lottery case:

Lottery meltdown
P1. The losing-claims in the trillion-ticket lottery are mutually inconsistent.
P2. If these losing-claims are mutually inconsistent, then certain privileged losing-claim judgments are mutually inconsistent.
P3. If these privileged losing-claim judgments are mutually inconsistent, then these privileged losing-claim judgments are unreliable.
P4. If these privileged losing-claim judgments are unreliable, then losing-claim judgments in lotteries more generally are unreliable.

C. Thus, losing-claim judgments in lotteries more generally are unreliable.

[11]Note, however, that even if there is indeed a rationale for singling out one proposition for rejection —that is, even if the proposition is based on different considerations and is also significantly less compelling than the others, considered individually—it does not follow that the *conjunction* of the others is more compelling. This will depend on further considerations. Still, the conjunction of the others *might* be more compelling, and so equal compellingness would forestall a potentially powerful objection (as we will see in Section 3).

If we understand epistemic privilege as scrutinized equal compellingness, then P2 will be true. But P3 will be false. We are still extremely reliable when it comes to picking losing tickets in a trillion-ticket lottery.[12] Thus, the step from *privileged inconsistency* to *privileged unreliability* is blocked. And this will be true also in more friendly lotteries, where the chances of winning are much higher than one in a trillion.

Still, the imagined arithmetical case does seem to present a significant skeptical threat, and so we should look to it for guidance. Why does the arithmetical case have skeptical import when the lottery case doesn't? More specifically, what relevant epistemic privilege, in addition to scrutinized equal compellingness, do the arithmetical judgments have that the lottery judgments lack?

Perhaps the difference is due to the number of propositions involved? The arithmetical case involved only two propositions while the lottery case involved huge numbers. I don't think the number of propositions involved is essential, however. Consider the following series:

$1 + 1 = 2$
$2 + 2 = 4$
…
$1,000,000 + 1,000,000 = 2,000,000$

If these propositions were shown to be inconsistent we would presumably still get the same meltdown effect, whereas the analogous lottery case poses no threat to our reliability at picking losing tickets. There is something else going on, but what?

I think the crucial difference is this: In the lottery case, no matter how many tickets are involved, the reasons for believing that any individual ticket will lose are *inconclusive*, and we know this beforehand. Thus, in forming the belief that a given ticket will lose, we know that we might be mistaken. By contrast, in the arithmetical case we are not aware of any such inconclusiveness. We take the reasons for believing that $2 + 2 = 4$ and so on to be utterly decisive. That is why we would be dumbfounded by the arithmetical inconsistency. We would be clueless as to why our intuitions have failed us, and so we would have no idea how to diagnose the situation. By contrast, in the lottery case we understand precisely why our judgment fails us in those rare winning-ticket cases, and so no threat is posed to our general reliability at picking losing tickets.

This provides a further constraint on the relevant notion of epistemic privilege: it should involve no inconclusiveness of the kind that renders the lottery case harmless. We have thus wound up with epistemic privilege as

---

[12]Other things in the vicinity might be threatened, however. For example, perhaps we cannot *know* that any given ticket will lose, even under the assumption that it will.

by-our-lights-conclusive scrutinized equal compellingness.[13] With these clarifications in hand, in the following section I will argue that the population ethical skeptical challenge fails.

3. *Why the skeptical challenge fails*

To make my critique more precise, let me first reconstruct the skeptical challenge along familiar lines, using the phrase 'POPULATION ETHICAL IMPOSSIBILITY RESULT' as a placeholder for some suitable impossibility theorem:

Moral meltdown

P1. POPULATION ETHICAL IMPOSSIBILITY RESULT
P2. If POPULATION ETHICAL IMPOSSIBILITY RESULT, then certain privileged population ethical intuitions are mutually inconsistent.
P3. If these privileged population ethical intuitions are mutually inconsistent, then these privileged population ethical intuitions are unreliable.
P4. If these privileged population ethical intuitions are unreliable, then moral intuitions are generally unreliable.

C. Thus, moral intuitions are generally unreliable.

The problems with this argument are twofold. First, no population ethical impossibility result (as far as I'm aware) meets the requirements concerning epistemic privilege (i.e., by-our-lights-conclusive scrutinized equal compellingness) set out in the previous section. If this is right then P2 is false. Second, the argument runs the risk of leaking into various other parts of philosophy, where impossibility results of various kinds are rather common (even though they are rarely highlighted with the kind of formal precision typical of population ethics). This point does not single out any particular premise as the culprit, but it provides a cautionary note with respect to the whole argument. Let us consider these two problems in turn.[14]

In Section 2, I presented the following adequacy condition from one of Arrhenius' impossibility theorems:

---

[13]Shouldn't we also build in something like a requirement of agreement among rational inquirers? I don't think that is necessary. Insofar as agreement among rational inquirers is lacking, we can instead take that to cast doubt on the idea that the proposition in question remains sufficiently compelling after due consideration and scrutiny. I will pursue this thought in Section 3.

[14]There is an additional way of challenging P2, namely, by questioning the alleged inconsistency of the conditions rather than their epistemic privilege. For example, Erik Carlson (2022) argues that the five conditions in Arrhenius' Sixth Theorem are inconsistent only given a certain technical assumption about the finite fine-grainedness of welfare levels, and he goes on to argue that this assumption should be rejected. Carlson also argues that this same critique applies to Arrhenius' five other impossibility theorems as well. See also Thomas (2018). I will not pursue this type of criticism, however.

The Non-Sadism Condition: An addition of any number of people with positive welfare is at least as good as an addition of any number of people with negative welfare, other things being equal.

I also said that this seems hard to deny. Indeed, I chose the Non-Sadism Condition for illustration precisely because of its apparent compellingness. The question, however, is whether there is an impossibility result whose conditions *together* meet the requisite by-our-lights-conclusive scrutinized equal compellingness. In what follows I will argue that there isn't. I cannot survey all of the impossibility results out there, of course. Instead I will focus on Arrhenius' *Third Impossibility Theorem*,[15] of which the Non-Sadism Condition is a part, and I will suggest that the problems I raise can be generalized.

According to Arrhenius' Third Theorem, '[t]here is no population axiology which satisfies the Egalitarian Dominance, the Inequality Aversion, the Non-Extreme Priority, the Non-Sadism, and the Quality Addition Condition'.[16] Here are the five conditions spelled out:

The Egalitarian Dominance Condition: If population A is a perfectly equal population of the same size as population B, and every person in A has higher welfare than every person in B, then A is better than B, other things being equal.

The Inequality Aversion Condition: For any triplet of welfare levels **A**, **B**, and **C**, **A** higher than **B** and **B** higher than **C**, and for any population A with welfare **A**, there is a larger population C with welfare **C** such that a perfectly equal population B of the same size as A∪C and with welfare **B** is at least as good as A∪C, other things being equal.

The Non-Extreme Priority Condition: There is a number *n* of lives such that for any population X, a population consisting of the X-lives, *n* lives with very high welfare, and a single life with slightly negative welfare, is at least as good as a population consisting of the X-lives and *n* + 1 lives with very low positive welfare, other things being equal.

The Non-Sadism Condition: An addition of any number of people with positive welfare is at least as good as an addition of any number of people with negative welfare, other things being equal.

The Quality Addition Condition: For any population X, there is a perfectly equal population with very high welfare such that its addition to X is at least as good as an addition of any population with very low positive welfare to X, other things being equal.[17]

The conditions might be evaluated piecemeal, by engaging in first-order population ethics. I will do that to some extent, restricting my attention to the Non-Sadism Condition and the Quality Addition Condition. First, however, I want to make a more general point.

---

[15]Arrhenius (forthcoming, § 11.8). Cf. Arrhenius (2000a, § 10.8). Arrhenius (forthcoming, p. 299) suggests that his Sixth Theorem is the most formidable one. But it involves more complicated conditions, and very much the same considerations apply. I will return to the Sixth Theorem in Section 4.
[16]Arrhenius (forthcoming, p. 317).
[17]Arrhenius (forthcoming, pp. 61, 145, 154, 95, 84–85).

Pacific Philosophical Quarterly published by University of Southern California and John Wiley & Sons Ltd.

Note that the conditions are all *principles*, not particular claims. Indeed, they aspire to hold for all *possible* populations, lives, additions and so on that meet the specifications.[18] Given this logical strength, it does not seem out of the question that an odd, completely unanticipated counterexample is lurking somewhere, rendering one or more of the conditions false but still *nearly true*. This would be analogous to how (long inaccessible) counterexamples to Newtonian mechanics were lurking, rendering the theory false and yet nearly true. Indeed, it is likely that *if* one of the conditions is false and yet as intuitively compelling as Arrhenius suggests, then this is precisely because the condition in question is nearly true.

I think we can draw two lessons from this nearly-true point. In the case of Newtonian mechanics, we should first of all not conclude that the kind of scientific reasoning based on empirical data that supported the theory is generally unreliable. It just went wrong in this particular case, precisely due to the logical strength of the theory's statements about the behavior of physical objects. Second, and more importantly, we should not conclude that our everyday judgments about the behavior of particular physical objects (about apples falling to the ground, or on one's head, for example) are generally unreliable. Analogously, if one of Arrhenius' principles is false but nearly true, we should first of all not conclude that the kind of reasoning based on moral intuitions that supported the principle is generally unreliable. Second, and more importantly, we should not conclude that our everyday moral intuitions about the social welfare of particular populations (let alone our moral intuitions in general) are unreliable.

Put in terms of the terminology from Section 2, the logical strength of the conditions leaves an inconclusiveness of the kind that allows our judgments about individual losing lottery tickets to retain reliability despite mutual inconsistency.[19]

Moving on to more specific considerations, note first that the Non-Sadism Condition is incompatible with *average utilitarianism*, according to which the best population is always the one with the highest average of individual welfare. (Adding a small number of people with negative welfare might lower the average less than adding a large number of people with lower-than-average positive welfare.) Average utilitarianism is of course a controversial view with seemingly indigestible implications,[20] but one might wonder whether it is so obviously false as

---

[18]Arrhenius (forthcoming, § 2.3). He nevertheless speaks of the intuitions in play as 'intuitions regarding particular cases' (forthcoming, p. 394).

[19]An analogous point could perhaps be made with respect to the imagined arithmetical and logical scenarios. If $2 + 2 = 4$ or *modus ponens* should turn out to be false, then this might be because they are necessary universal generalizations with a few extremely odd counterexamples, rendering them false but nearly true. And if not even these scenarios would pose a skeptical threat, then it is hard to believe that population ethical impossibility results do.

[20]See, for example, Carlson (2017, pp. 12–13) and Arrhenius (forthcoming, pp. 82–83).

to not jeopardize the by-our-lights-conclusive scrutinized compellingness of the Non-Sadism Condition.[21]

A further and more serious problem is that not all of Arrhenius' adequacy conditions have the same air of compellingness as the Non-Sadism Condition. Consider the Quality Addition Condition:

The Quality Addition Condition: For any population X, there is a perfectly equal population with very high welfare such that its addition to X is at least as good as an addition of any population with very low positive welfare to X, other things being equal.

(The phrases 'very high welfare' and 'very low positive welfare' should here be understood as signifying levels of individual, not total, welfare.)

The Quality Addition Condition may seem plausible. Indeed, the condition gains support from the kind of intuitions that led Parfit to famously choose the label *The Repugnant Conclusion* for what is essentially a denial of the Quality Addition Condition.[22] Still, the condition is controversial, and much more so than the Non-Sadism Condition. Indeed, the Quality Addition Condition is controversial in a way that is hard to reconcile with the idea that it enjoys the kind of epistemic privilege required to sustain the skeptical challenge. For example, the condition is incompatible with *total utilitarianism*, according to which only the total *quantity* of individual welfare matters for social welfare. Similarly, the condition rules out *prioritarian* modifications of total utilitarianism, according to which only the total quantity of individual welfare, weighted by a strictly concave function, matters for social welfare. These views may be wrong, but they are not obviously wrong, and they have quite a few adherents in contemporary population ethics.[23]

Thus, the Quality Addition Condition does seem to involve an inconclusiveness of the relevant kind. Even if we find the condition highly plausible or even compelling, we know beforehand that the condition might be mistaken. We would not find ourselves morally dumbfounded should the condition turn out to be false (in the way that we would find ourselves arithmetically dumbfounded if 2 + 2 = 4 turned out to be false). We would just conclude that our intuitions in this particular case are off track, perhaps because we have difficulties in imagining large quantities, or because we have a distorted conception of what having low positive welfare is like.[24] Put in terms of my earlier terminology, the Quality Addition Condition does

---

[21]For a recent defense of average utilitarianism, see Pressman (2015).

[22]Parfit (1984, p. 388).

[23]Cf. Zuber et al. (2021), in which no less than 29 authors jointly declare that 'avoiding the Repugnant Conclusion is not a *necessary* condition for a minimally adequate candidate axiology, social ordering, or approach to population ethics' (p. 380, emphasis in original). This amounts to a rejection of the Quality Addition Condition.

[24]See, for example, Tännsjö (2002) and Huemer (2008). Cf. Zuber et al. (2021, p. 380).

not fulfill the requirement of by-our-lights-conclusive scrutinized compellingness.

Moreover, as the Quality Addition Condition is incompatible with the conjunction of the four other conditions in the Third Theorem, these other conditions can be invoked to throw further doubt on the Quality Addition Condition. Thus, we can formulate a valid argument as follows:

P1. The Egalitarian Dominance Condition
P2. The Inequality Aversion Condition
P3. The Non-Extreme Priority Condition
P4. The Non-Sadism Condition

C. ¬ The Quality Addition Condition

Assuming that these premises are jointly sufficiently compelling, we get further grounds for questioning the supposed epistemic privilege of the Quality Addition Condition.[25]

It might be objected that analogous arguments could be made against *each* of the five conditions, and so these arguments are mutually destructive and thus irrelevant to the issue at hand. However, this objection presupposes that the conditions are all *equally* compelling, and this might not be the case (recall the requirement of by-our-lights-conclusive scrutinized *equal* compellingness). I have indicated why I think the Quality Addition Condition is at least less compelling than the Non-Sadism Condition, and it may just be that even the conjunction of the four other conditions is more compelling than the Quality Addition Condition on its own. I will not be able to settle this issue here, of course. The point is just that it presents a further obstacle that a successful skeptical challenge based on Arrhenius' Third Theorem would need to overcome.

I conclude that at least one of the conditions in Arrhenius' Third Theorem is questionable enough that it cannot sustain a skeptical challenge. Note, however, that even in the absence of any specific considerations threatening one or more of the conditions, we have still been given no good reason to distrust our moral intuitions about the social welfare of particular populations (perhaps comparing the population of Sweden today with the one during the Black Death of 1350). This is due to the logical strength of the conditions, being principles rather than particular claims. In other words, even if we don't know which condition to reject, it does not follow that our population ethical intuitions (let alone our other moral intuitions) are unreliable across the board.[26]

---

[25]Cf. Zuber et al. (2021, pp. 380–381).
[26]It is worth noting here that according to *moral particularism*, our moral intuitions are fundamentally and thoroughly context-sensitive and cannot be accommodated by *any* moral principles, no matter how plausible they may appear at first glance (see, e.g., Dancy, 2017). Thus, from a particularist point of view, impossibility results à la Arrhenius are just what we should expect.

Arrhenius' Third Theorem is just one example, but it is illustrative. There is to my knowledge no population ethical impossibility result whose adequacy conditions enjoy the requisite epistemic privilege.[27] If this is right, then no skeptical conclusion is called for.

But suppose this isn't right. Suppose (for *reductio*) that the population ethical skeptical argument does manage to establish its conclusion. If so, then the worry is that there will be a host of analogous, and equally successful, skeptical arguments throughout philosophy, forcing us into a much broader philosophical skepticism.[28] It is a familiar lesson from philosophy in general, after all, that our intuitions often pull in different directions, and consequently that various seemingly plausible or even compelling claims often cannot be reconciled.

This point is rather obvious but no less important for that. It can be given a particularly sharp illustration by considering the *free-will problem*, as presented by Peter van Inwagen.[29] The way van Inwagen sets up the problem is similar to how Arrhenius and others have set up population ethical impossibility results. In van Inwagen's presentation, the following five propositions correspond to population ethical adequacy conditions, and their mutual inconsistency is a matter of straightforward logic:

(1)   Free will is incompatible with determinism.
(2)   Free will is incompatible with indeterminism.
(3)   If free will is incompatible both with determinism and with indeterminism, then free will does not exist.
(4)   If free will does not exist, then moral responsibility does not exist.
(5)   Moral responsibility does exist.

'Free will' is here understood as the *ability to do otherwise*. More precisely:

The free-will thesis is the thesis that we are sometimes in the following position with respect to a contemplated future act: we simultaneously have both the following abilities: the ability to perform that act and the ability to refrain from performing that act.[30]

Concerning (1), it is indeed hard to see how free will in this sense could be compatible with determinism. If every act we perform is

(a)   a consequence of
(b)   the distant past and
(c)   the laws of nature,

---

[27]For example, very much the same considerations apply to Arrhenius' five other impossibility theorems (forthcoming, ch. 11).
[28]Arrhenius (forthcoming, p. 393) anticipates this worry but does not pursue it.
[29]van Inwagen (2008, pp. 327–328).
[30]van Inwagen (2008, p. 329).

and because all of (a)–(c) are surely beyond our control, then how could we ever have the ability to do otherwise? Concerning (2), it is hard to see how introducing an element of pure chance into the processes by which our actions come about is supposed to help. (3) is trivial, and (4) seems obvious as it would surely be grossly unfair to hold someone accountable for some act that they had no ability to refrain from. (5) seems equally obvious, as it was surely not unfair to hold Ted Bundy accountable for his killing spree, for example.

All of this is controversial, of course, and I have only sketched the contours of the problem. Still, each of (1)–(5) seems compelling on the face of it, and I submit that the conditions for epistemic privilege are fulfilled at least to the same degree as in any extant population ethical impossibility result. The age-old controversy around (1), for example, is no different in this respect than the controversy around the Quality Addition Condition.

In presenting a set of irreconcilable and yet individually plausible or even compelling propositions, the free-will problem is a typical philosophical problem, just like the mind–body problem, the problem of universals, the problem of induction, the problem of the many, the problem of personal identity across time, and so on. To this list we can add the problem of population ethics.

Setting aside worries about self-defeat, a broader philosophical skepticism is at least a lot harder to swallow than a local moral skepticism. Perhaps more interestingly, however, we don't in fact see philosophers in other areas working under a specter of skepticism. For example, while proposition (4) in the free-will problem—that is, that moral responsibility requires the ability to do otherwise—may seem completely obvious, it was famously challenged by Harry Frankfurt, using an ingenious counterexample.[31] The merits of Frankfurt's counterexample are controversial, of course, but the point is that it wasn't put forward in an anxious attempt to avoid skepticism about the issues at hand.

This puts population ethical impossibility results into perspective. I think there is a tendency within philosophical ethics to expect too much from the discipline, and skeptical or anti-realist conclusions are often drawn from arguments whose analogues do not look promising in most other areas of philosophy. I'm thinking especially of arguments from moral disagreement, which in various versions have played a central role in recent metaethics,[32] and the skeptical and anti-realist arguments from population ethical impossibility results are in the same vein. The underlying idea seems to be that ethics is somehow supposed to be devoid of the kind of deep inter- and intrapersonal conflicts of intuition that are characteristic of most philosophical inquiry. Perhaps there is something to this idea, but it would in any case need to be defended as a part of the population ethical skeptical challenge.

[31]Frankfurt (1969). For discussion, see Robb (2020).
[32]For an extensive critical survey of such arguments, see Enoch (2011, ch. 8).

4. *The anti-realist challenge and why it too fails*

The population ethical anti-realist challenge faces similar problems, and so we already have some resources on the table for dealing with it. However, there are a few different versions of the anti-realist challenge, and they require somewhat different treatment. As I mentioned in Section 1, there is the McMahan-Kagan-Temkin challenge, according to which population ethical impossibility results pose a threat to 'realism in ethics'. We might call this:

(i) the *generic* challenge.[33]

In addition, there are more specific anti-realist challenges:

(ii) the *error theorist* challenge, according to which moral judgments are shown to be systematically false[34]; and

(iii) the *non-cognitivist* challenge, according to which moral judgments are shown to be emotions or conative attitudes rather than beliefs.[35]

I will not discuss the generic challenge, as it is not independent of the more specific ones. In the rest of this section, I will explain and criticize these challenges, beginning with the error theorist one.

As a first component of the error theorist challenge, Christopher Cowie appeals to the so-called 'fixed points' approach to moral concepts, proposed by Terence Cuneo and Russ Shafer-Landau.[36] On this approach, a subset of especially compelling, but still substantive, moral propositions are held to be *conceptual* truths. Here are some of Cuneo and Shafer-Landau's suggested examples:

- It is pro tanto wrong to engage in the recreational slaughter of a fellow person.
- It is pro tanto wrong to break a promise on which another is relying simply for convenience's sake.
- It is pro tanto wrong to humiliate others simply for pleasure.
- It is pro tanto wrong to torture others just because they have inconvenienced you.
- It is pro tanto wrong to impose severe burdens on others simply because of their physical appearance.[37]

---

[33]McMahan (2013, p. 34); Temkin (2012, p. 521, n. 42). See also Fleurbaey et al. (2009), pp. 274, 284), who put the challenge in terms of 'avoiding moral nihilism'.

[34]The error-theorist challenge is developed by Cowie (2022, § 3).

[35]The non-cognitivist challenge is suggested by Arrhenius (forthcoming, pp. 397–398). There are hints at a non-cognitivist challenge also in Temkin's Kagan-quote (Temkin, 2012, p. 521, n. 42).

[36]Cuneo and Shafer-Landau (2014, § 2).

[37]Cuneo and Shafer-Landau (2014, p. 405).

These propositions, or propositions like these, are the moral fixed points. The idea is that anyone who is competent with the concept of moral wrongness is thereby in a position to realize, for example, that it is pro tanto wrong to humiliate others simply for pleasure. Likewise for the other central moral concepts.

As a second component, Cowie takes on board Arrhenius' idea that the conditions involved in the impossibility theorems are intuitively highly compelling (Cowie focuses on Arrhenius' Sixth Theorem). Indeed, the idea is that the conditions are compelling enough to count as conceptual truths on the fixed-points approach. To support this claim, Cowie appeals to the four 'marks' of conceptual truths suggested by Cuneo and Shafer-Landau:

First, p is, if true, necessarily true. […] Second, p enjoys framework status, fixing the boundaries as to what counts as a type of subject matter. […] Third, p's denial would tend to evoke bewilderment among those competent with its constituent concepts […]. And, fourth, p is knowable a priori, simply by adequately understanding its constituent concepts and their relations to one another.[38]

Cowie suggests that all of the conditions involved in Arrhenius' Sixth Theorem display these marks.[39]

The upshot is that Arrhenius' Sixth Theorem reveals conceptual incoherence with respect to the concept of goodness, or intrinsic value. The concept of intrinsic value is thus like the concept of a round square, and so nothing could fall under it. And assuming that this concept is sufficiently central to moral thought and discourse, the incoherence will render all moral propositions false. Thus, we end up with a moral error theory.[40]

Just like the skeptical challenge promised to sidestep various contested issues in moral epistemology, the error theorist challenge promises to sidestep various contested issues in moral metaphysics. If it turns out that moral thought and talk is like round-square thought and talk, then we can set aside controversial issues about the putative metaphysical weirdness of moral properties and facts, for example.[41]

The error theorist challenge does not succeed, however. An immediate problem is that the whole point of Arrhenius' theorems is to show that the conditions are mutually *inconsistent*. And if they are indeed mutually inconsistent, then they cannot all be conceptual *truths*. Moreover, according to Cowie's own error theorist conclusion, the conditions are all *false*. And if so, again, they cannot be conceptual truths. Cowie does address this problem, however, at least indirectly. In answering a general objection to the

---

[38]Cuneo and Shafer-Landau (2014, pp. 407–408).

[39]Cowie (2022, p. 285).

[40]Cowie (2022, pp. 280–281) acknowledges, however, that we might rather end up with a more limited, merely axiological error theory, depending on the importance of axiology to moral thought and discourse.

[41]Cowie (2022, p. 277).

fixed-points approach, he suggests that the conceptual truths in question 'should be recast as conditionals with "*if anything is wrong/bad*" as the antecedent and substantive moral propositions as the consequents'.[42] This solves the problem at hand, as the resulting conditionals will not be inconsistent and the error theory will not render them false. Another way of making essentially the same modification is to talk, as I will do below, in terms of conceptual *commitments* rather than conceptual truths.

Three problems remain, however. First, the challenge might be countered by questioning the merits of the fixed-points approach. One can object on Moorean open-question grounds that 'thin' moral concepts such as goodness or rightness are devoid of descriptive content, or at least that they are devoid of *substantive* descriptive content, which goes beyond formalities such as universalizability and supervenience. This objection to the fixed-points approach becomes especially pressing in light of the error theory's implication, for example, that recreational torture is not even pro tanto wrong. Assuming that the other components of the error theorist challenge stand up to scrutiny, its proponents need to explain why we should stick with the controversial and highly theoretical fixed-points approach rather than with our most basic moral intuitions or convictions (especially if these are compelling enough to count as conceptual commitments, as the error theorist challenge would have it).

Second—and this relates very much to my earlier discussion of the skeptical challenge—Arrhenius' conditions are just not suitable candidates for being conceptual commitments. Or, at least not all of them are. Consider the following two conditions from Arrhenius' Sixth Theorem:

The Weak Non-Sadism Condition: There is a negative welfare level and a number of lives at this level such that an addition of any number of people with positive welfare is at least as good as an addition of the lives with negative welfare, other things being equal.

The Weak Quality Addition Condition: For any population X, there is a perfectly equal population with very high positive welfare, and a very negative welfare level, and a number of lives at this level, such that the addition of the high welfare population to X is at least as good as the addition of any population consisting of the lives with negative welfare and any number of lives with very low positive welfare to X, other things being equal.[43]

Just like the Non-Sadism Condition and the Quality Addition Condition considered in Section 3, the Weak Non-Sadism Condition and the Weak Quality Addition Condition are incompatible with average and total utilitarianism, respectively. And it seems far-fetched to suppose that the falsity of these views can be ascertained on conceptual grounds alone. This point pertains to Cuneo and Shafer-Landau's third mark of conceptual

---

[42]Cowie (2022, pp. 286–287).
[43]Arrhenius (forthcoming, pp. 96, 86).

commitments, namely that 'p's denial would tend to evoke bewilderment among those competent with its constituent concepts—a response to the effect that its denial would be almost crazy'.[44] Even if you find average and total utilitarianism highly implausible, it would be an exaggeration to suggest that they evoke bewilderment, or that their proponents are almost crazy.

Third, insofar as the error theorist challenge does succeed, it is not clear why analogous strategies would not wreak havoc throughout philosophy. For example, the five propositions that constitute the free-will problem (from Section 3) would seem to be just as good candidates for conceptual-commitment status as Arrhenius' conditions. Thus, just like in the case of the skeptical challenge, opponents of the error theorist challenge can appeal to philosophy in general as a partner in guilt.

Let us move on to the non-cognitivist challenge. Why would population ethical impossibility results speak in favor of non-cognitivism? One reason, suggested by Arrhenius, is that non-cognitivist views will have an easier time accommodating *inconsistency* among our moral intuitions or judgments. On cognitivist views, moral inconsistency will amount to a straightforward logical error. But if our moral judgments are emotions or conative attitudes rather than beliefs, then presumably they cannot be *logically* inconsistent anyway. Inconsistency on non-cognitivist views will rather have to be explicated in practical terms, for example in terms of unrealizable action-guidance. But because the inconsistencies putatively revealed by population ethical impossibility results are so remote from the concerns of real life, they will have little or no practical import. Hence, given non-cognitivism we might be, as Arrhenius puts it, 'justified in worrying less' about these inconsistencies.[45]

This is not a very good argument, however. First, it has an air of wishful thinking about it. Sure, it would be nice if we needed to worry less about the inconsistencies in question, but how does that translate into a reason to accept non-cognitivism? If we had *independent* support for the claim that we need to worry less, then non-cognitivism might gain support from making the right prediction, so to speak. As it stands, however, the argument seems at best sufficient to motivate the non-cognitivist solution, not to justify it.[46]

Second, and as we have already noted, philosophy is rife with apparent inconsistencies, and usually ones that are very remote from the practical concerns of real life. For example, while thought experiments involving teleportation bring out deep tensions within our everyday conception of personal identity across time, these tensions don't present us with difficulties in keeping track of who's who in real life. Similarly, there are deep puzzles

---

[44]Cuneo and Shafer-Landau (2014, pp. 407–408).
[45]Arrhenius (forthcoming, p. 398).
[46]In fairness to Arrhenius, he may not be after more than to motivate the non-cognitivist solution.

about material objects, such as the ship of Theseus and the problem of the many, but these don't obstruct our everyday navigation of the physical world. Thus, insofar as population ethical impossibility results support non-cognitivism about moral thought and discourse, philosophical puzzles of various kinds are bound to support non-cognitivism about most of philosophy as well. To be fair, it may be reasonable to suppose that there is a weak pro tanto reason of this kind, which is overridden by considerations supporting cognitivism about philosophical inquiry in general. But if we take the reason to be a strong or overriding one, we will end up with a very radical conclusion indeed.

5.   *Conclusion*

I have argued that the skeptical and anti-realist challenges based on population ethical impossibility results both fail.

Concerning the skeptical challenge, there is no population ethical impossibility result whose adequacy conditions fulfill the requisite epistemic privilege (at least as far as I'm aware). Moreover, even if there is such a result, it will just be yet another philosophical puzzle, and any skeptical conclusions are bound to generalize. This is not to deny, however, that a certain limited skepticism is apt with respect to some or even most philosophical issues. For the most part, we cannot reasonably have the same degree of confidence in our philosophical positions that we have in well-established scientific theories, for example.[47] But philosophical intuitions being unreliable across the board is another matter.

Just like the skeptical challenge, both versions of the anti-realist challenge would seem to spread malignantly to philosophy in general. Moreover, both versions face additional problems. Concerning the error theorist challenge, the fixed-points approach to moral concepts is a controversial piece of philosophical theorizing, and as such it can hardly sustain the error theorist conclusion, for example, that being tortured is not even pro tanto bad for the victim. And even granting the fixed-points approach, at least some of Arrhenius' conditions are unsuitable as candidate conceptual commitments. Similarly, while non-cognitivism about moral thought and language may permit us to worry less about putative population ethical inconsistencies, we would need independent justification for granting ourselves such permission.[48]

Department of Historical, Philosophical and Religious Studies
Umeå University

REFERENCES

Arrhenius, G. (2000a). *Future Generations: A Challenge for Moral Theory*. Dissertation. Uppsala University.

Arrhenius, G. (2000b). 'An Impossibility Theorem for Welfarist Axiologies,' *Economics & Philosophy* 16(2), pp. 247–266.

Arrhenius, G. (2004). 'The Paradoxes of Future Generations and Normative Theory,' in J. Ryberg and T. Tännsjö (eds) *The Repugnant Conclusion*. Dordrecht: Kluwer, pp. 201–218.

Arrhenius, G. (2011). 'The Impossibility of a Satisfactory Population Ethics,' in H. Colonius and E. Dzhafarov (eds) *Descriptive and Normative Approaches to Human Behavior*. Singapore: World Scientific Publishing Company, pp. 1–26.

Arrhenius, G. (forthcoming). *Population Ethics: The Challenge of Future Generations*. Oxford: Oxford University Press (Manuscript dated June 2018.).

Blackorby, C. and Donaldson, D. (1991). 'Normative Population Theory: A Comment,' *Social Choice and Welfare* 8(3), pp. 261–267.

Carlson, E. (1998). 'Mere Addition and Two Trilemmas of Population Ethics,' *Economics & Philosophy* 14(2), pp. 283–306.

Carlson, E. (2017). 'Aggregate and Average Utilitarianism,' in J. E. Crimmins (ed.) *The Bloomsbury Encyclopedia of Utilitarianism*. London and New York: Bloomsbury Publishing, pp. 11–13.

Carlson, E. (2022). 'On Some Impossibility Theorems in Population Ethics,' in G. Arrhenius, K. Bykvist, T. Campbell and E. Finneron-Burns (eds) *The Oxford Handbook of Population Ethics*. Oxford and New York: Oxford University Press, pp. 204–224.

Cowie, C. (2022). 'A New Argument for Moral Error Theory,' *Noûs* 56(2), pp. 276–294.

Cuneo, T. and Shafer-Landau, R. (2014). 'The Moral Fixed Points: New Directions for Moral Nonnaturalism,' *Philosophical Studies* 171(3), pp. 399–443.

Dancy, J. (2017). 'Moral Particularism,' in E. N. Zalta (ed.) *The Stanford Encyclopedia of Philosophy*, Winter 2017 edn. <https://plato.stanford.edu/archives/win2017/entries/moral-particularism/>.

Enoch, D. (2011). *Taking Morality Seriously: A Defense of Robust Realism*. Oxford: Oxford University Press.

Fleurbaey, M., Tungodden, B. and Vallentyne, P. (2009). 'On the Possibility of Nonaggregative Priority for the Worst Off,' *Soc Philos Policy* 26(1), pp. 258–285.

Frankfurt, H. G. (1969). 'Alternate Possibilities and Moral Responsibility,' *The Journal of Philosophy* 66(23), pp. 829–839 Reprinted in the Importance of What We Care About: Philosophical Essays, Cambridge University Press, 1988.

Huemer, M. (2008). 'In Defence of Repugnance,' *Mind* 117(468), pp. 899–933.

van Inwagen, P. (2008). 'How to Think About the Problem of Free Will,' *The Journal of Ethics* 12(3/4), pp. 327–341.

van Inwagen, P. (2015). *Metaphysics*, Fourth edn. Boulder, CO: Westview Press.

Joyce, R. (2006). *The Evolution of Morality*. Cambridge, MA: The MIT Press.

McMahan, J. (2013). 'Causing People to Exist and Saving People's Lives,' *The Journal of Ethics* 17(1/2), pp. 5–35.

Ng, Y.-K. (1989). 'What Should We Do About Future Generations? Impossibility of Parfit's Theory X,' *Economics & Philosophy* 5(2), pp. 235–253.

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

Pressman, M. (2015). 'A Defence of Average Utilitarianism,' *Utilitas* 27(4), pp. 389–424.

Robb, D. (2020). 'Moral Responsibility and the Principle of Alternative Possibilities,' in E. N. Zalta (ed.) *The Stanford Encyclopedia of Philosophy*, Fall 2020 edn. <https://plato.stanford.edu/archives/fall2020/entries/alternative-possibilities/>.

Tännsjö, T. (2002). 'Why we Ought to Accept the Repugnant Conclusion,' *Utilitas* 14(3), pp. 339–359.

Temkin, L. S. (2012). *Rethinking the Good: Moral Ideals and the Nature of Practical Reasoning*. Oxford: Oxford University Press.

Thomas, T. (2018). 'Some Possibilities in Population Axiology,' *Mind* 127(507), pp. 807–832.

Zuber, S. et al. (2021). 'What Should We Agree on About the Repugnant Conclusion?' *Utilitas* 33(4), pp. 379–383.