# The Possibility of Morality

Philip Brown
philip.brown@postgrad.manchester.ac.uk
Penultimate draft; forthcoming in *Philosophical Studies.*
(Please cite published version.)

One of the central debates in metaethics is over whether or not there are moral facts. A related but neglected question is: *could* there be moral facts? That is to say, could states of affairs or characters or institutions be morally good or bad; could actions be morally permissible, obligatory, or forbidden? (Throughout I will be concerned with metaphysical modalities.)

Naturally, moral realists will respond that moral facts are indeed possible, because many such facts actually exist, and actuality entails possibility.

Noncognitivist theories will give various answers to the question. For instance, emotivists, such as A. J. Ayer, claimed moral utterances serve to express feelings and not to report facts (Ayer 1936). Emotivists would presumably have held that, once this is understood, there is no sense to be made of the question of the possibility of moral facts, because moral language just lacks a representational function. By contrast, recent developments in noncognitivism have been focussed on attempting to show how we can retain all the realist-sounding trappings of moral language without committing to an external moral reality, in part by adopting deflationary accounts of truth, propositions, and properties.[1] If successful, this strategy would permit them to talk about moral facts, both actual and possible, without committing them to a moral reality that exists independently of moral discourse and practice.

Moral error theorists, on the other hand, side with the realists on the ontological commitments of moral language. A sentence such as "Stealing is wrong" does not primarily function to express the speaker's attitude to stealing, according to both realists and error theorists, but to assert that acts of stealing have the property of wrongness independently of anyone‟s actual attitude to stealing. Error theory differs from realism, however, in its claim that no statement attributing a moral property to an object is true, because moral properties do not exist.[2] But non-existence does not entail impossibility. So the question I want to focus on in this paper is: What should error theorists say about the possibility of moral facts?

Despite the neglect of this question, we can find resources for answering it by examining some of the different versions of error theory that have emerged over the years. The focus of my discussion will be on what error theorists should say about the following claims:

> *Weak:* There is a possible world *W* where there are moral facts.
>
> *Strong:* There is a possible world *W*, identical to the actual world in all non-moral respects, where there are moral facts.

Anyone who believes in the existence of moral facts at this world should accept both *Strong* and *Weak. Strong* is a trivial entailment of the proposition that there exist moral facts at the actual world, because the actual world is identical to the actual world in all non-moral respects. And *Strong* entails *Weak.* The interesting questions are: What should someone who

[1] See Dreier (2004) for a review.
[2] The *locus classicus* for moral error theory is Mackie (1977).

denies the existence of moral facts at this world say about *Strong* and *Weak*, to remain consistent with their denial of moral facts at this world? And what arguments can they give in favour of their positions on *Strong* and *Weak*?

In the first section, I show that every version of the error theory entails the negation of *Strong*, while at least two versions entail the negation of *Weak.* This may seem to make available a modal counterargument to error theory, which takes the possibility of moral facts as a premise. However, I show in section 2 that any such argument would be question begging. And this will allow us to conclude that moral error theorists are in a stronger position than has been previously noticed: able to deny not just the actuality, but the possibility of moral facts, without making them any more vulnerable to their opponents.

## 1. The modal status of error theory

J.L. Mackie provided the two best known arguments for moral error theory: the *argument from relativity* and the *argument from queerness* (Mackie 1977, chapter 1). The kernel of the first argument is that the observed diversity in moral opinions at different times and places is "more readily explained by the hypothesis that they reflect ways of life than by the hypothesis that they express perceptions, most of them seriously inadequate and badly distorted, of objective values" (ibid.: 37). Assume that Mackie is right both about the diversity of moral opinions at different times and places and about the best explanation of this fact. Then any world that is a non-moral duplicate of our own will be one at which there is a diversity of moral opinions that is best explained in the way that Mackie explains it. Unless we think we should believe some explanation for the diversity of moral opinions other than the best one, then we should conclude that *Strong* is false. The argument from relativity is toothless against *Weak*, however. For we can imagine worlds where there is a complete consensus over moral matters; and we can imagine worlds where the diversity of moral opinions is best explained by some other factor, such as differences in the reasoning capacities of individuals or differences in their access to empirical information that bears on moral matters. Error theory might still be true at such worlds, but the argument from relativity would not establish it.

If the argument from queerness is sound, it also refutes *Strong*. Moral properties are queer, for Mackie, because if they existed they would be objective (in the sense that their instantiation would be independent of anyone's opinions, feelings or decisions) and prescriptive (in the sense that knowing that something is, for instance, wrong, tells you that you shouldn't do it). This creates a metaphysical worry, because it doesn't look as if such properties fit into a naturalistic world-view, and an epistemological worry in that being nonnatural (and thus acausal) it is mysterious how we could ever have knowledge of them (Mackie 1977, chapter 1). In light of this, Mackie thought it more reasonable to accept a debunking explanation of moral discourse.[3] If Mackie is right in his analysis of what moral terms purport to refer to, and if he is right that it is most reasonable to accept a debunking explanation of our use of moral language in the actual world, then it is most reasonable to

[3] To "de-bunk" is to remove the nonsense (bunk) from. In this case, the nonsense is objective moral values. The argument from relativity is one example of a debunking argument. Another of Mackie's is based around his belief in projectivism, which he takes from Hume. According to Mackie, that which we desire is seen as good because we desire it; but we mistakenly reverse the direction of dependence, thinking that we desire what we desire because it *is* good (Mackie 1977: 43). More recently, moral sceptics have focussed on evolutionary debunking arguments, which try to account for moral behaviour by showing how it enhanced the reproductive fitness of our ancestors. Very roughly, our moral attitudes do not correspond to moral facts, on this view, but are caused (at least in part) by features of our psychology that have been shaped by our evolutionary history; so we can have no confidence that our moral beliefs are true. See Joyce (2006) for an argument of this type.

accept a debunking explanation of our counterparts' use of moral language in any non-moral duplicate of our world. After all, the conditions that make the debunking explanation most reasonable to accept in our world will be present in the non-moral duplicate of our world. Does the argument from queerness, if sound, refute *Weak*? It doesn't look like it. We might, for instance, be the most diehard naturalists about the ontology of this world, but be prepared to grant there are possible worlds with non-naturalistic ontologies. Or (perhaps relatedly) we might be able to conceive of worlds at which there are no plausible debunking explanations for the use of moral terms at that world, and so the most reasonable thing to say about moral terms there is that they successfully refer. Certainly these possibilities are not ruled out by anything contained within the argument from queerness, though a defender of the argument might want to provide supplementary arguments to rule out such possibilities and refute *Weak*.

Bart Streumer's argument for error theory shares some similarities with the argument from relativity. He begins by defending Frank Jackson's (1998) argument that normative properties are identical to descriptive properties (to be precise, Streumer defends a conditional form of Jackson's conclusion: *if there are normative properties*, then they are identical to descriptive properties) (Streumer 2008). In a subsequent paper he argues that if normative properties are identical to descriptive properties this entails that it is possible (in principle) to say which descriptive properties they are identical to (Streumer 2010). But, according to Streumer (2010), it is not possible (even in principle) to say which descriptive properties normative properties are identical to, because this is determined by how the folk apply normative terms and the folk don't agree about how to apply normative terms. From these premises it follows that there are no normative properties at this world.[4] Assuming that Streumer's premises are true, any world that is non-morally identical to ours will be a world at which it is not possible to say which descriptive properties normative properties are identical to, because the folk at that world will not agree in their usage of normative terms either. So Streumer's argument rules out *Strong*. As I have sketched his thesis, it would not rule out *Weak*, because we can imagine a world where the folk *do* agree about how to apply normative terms. However, on Streumer's view, this sort of agreement is not sufficient, because future members of the linguistic community might grow up to disagree with their parents on how to apply normative terms, and thus the normative beliefs of the parents and children would form an inconsistent set (ibid.). So Streumer's thesis seems to rule out *Weak* as well as *Strong*.

Richard Joyce's argument for error theory also rules out *Strong*. In *The Myth of Morality* Joyce argues that any reason that an agent has for acting in one way rather than another must be ultimately grounded in her desires or motivations (so all such reasons are 'internal' in the sense made famous by Bernard Williams (1981)). But putative moral reasons are not supposed to be grounded in the desires or motivations of the agents to whom they are attributed: we wouldn't excuse a murderous tyrant from moral criticism if the evidence suggested that, of the choices available to her, the ones she made best satisfied her desires. So, Joyce concludes, putative moral reasons are not real reasons for acting. There would be reasons that applied to all agents, regardless of their particular desires or motivations, if there were convergence in their desires and motivations. Although there is not actually such

[4] Streumer's thesis has the interesting consequence that if it is true we can have no reason to believe it. Indeed a number of philosophers have argued that Mackie"s version of the error theory also has this consequence, on the grounds that epistemic facts, if they exist, are no less categorical or objective than moral facts, and thus if we reject the latter we ought to reject the former as well. This is taken to be a *reductio* moral error theory. See Cuneo (2007) for a book-length version of this argument. Even if this objection were fatal to views like Mackie's and Streumer's, they may leave views like Joyce's and Loeb's (which we turn to presently) untouched.

convergence, Joyce seems to consider that it would be sufficient if the convergence came about under ideal circumstances (e.g. of full knowledge, full imaginative acquaintance with relevant possibilities, flawless deliberation). But Joyce argues that there is no reason to expect that such a convergence would take place, even in ideal circumstances, because there is no principled reason for thinking that an agent doing something that most of us would find morally abhorrent must be making an error in reasoning or be lacking some vital information (Joyce 2001, chapters 3 and 4).

Joyce's argument does not seem to rule out *Weak*. Even if it is true that there is no reason to expect a convergence in the motivations of agents *in this world* under ideal circumstances, this is a contingent fact about the group of agents in this world, rather than a necessary fact about groups of agents. In a world where all agents care equally for the wellbeing of themselves and all other agents, and have an equal regard for all non-agents, there plausibly would be such a convergence.

Lastly, Don Loeb has argued that the persistence of fundamental disputes between cognitivists and noncognitivists "is evidence that inconsistent elements – in particular, commitments both to and against objectivity – may be part of any accurate understanding of the central moral terms" (Loeb 2008: 357-58). He adds "If … moral vocabulary is best understood as semantically incoherent, the metaphysical implication is with respect to that vocabulary there is nothing in particular to be a realist *about* – no properties, that is, that count as referents of the moral terms" (ibid.). If Loeb is right, then moral terms do not even potentially refer to anything – at this world or any other – so there cannot be a world at which there are moral facts. Loeb's argument, if sound, would show that all propositions attributing moral properties are necessarily untrue, and thus that error theory is necessarily true (Streumer's thesis also entails this).

What has been shown is that although error theorists have hitherto only denied there are moral facts at this world, their arguments support the stronger conclusion that there are no moral facts at any world that is non-morally identical to ours, and, in the case of Streumer and Loeb, at any possible world. But we often find that the stronger the claim we make, the more vulnerable it is to counterargument. In particular, strengthening error theory in the way I have done may make it look vulnerable to a modal counterargument. In the next section I examine modal arguments against error theory, and show that the strengthening of the error theorist's claim does not expose it in this way.

## 2. A modal argument against error theory

Given the conclusion of the previous section, it becomes clear that moral realists need only show that moral facts are possible, in order to refute error theory. This, however, will not be so simple a matter as declaring that we can imagine a world at which there are moral facts. Doing that will only lead people to imagine worlds much like our own, and error theorists have arguments to the effect that there are no moral facts at those worlds.

It's true that if we could establish *Weak* then we would have all that we needed to refute any version of the error theory that entailed that error theory is necessarily true. However, if this entailment is backed up by argument, then any attempt to refute the error theory by appeal to *Weak* will be question-begging if it merely asserts *Weak* without addressing the error theorist's argument to the contrary.

Versions of the error theory that do not entail that error theory is necessarily true are under no threat from *Weak*. But they may appear to be threatened by *Strong*, particularly if one of the most widely accepted metaethical theses is true:

> *Supervenience:* moral properties (if they exist) supervene on non-moral properties, such that the subvening non-moral properties wholly determine the supervening moral properties.[5]

If we could establish that there is a possible world *W*, non-morally identical to our own, at which there are moral facts, then given *Supervenience*, this would show that there are moral facts at our world. For if the non-moral facts at *W* determine a set of moral facts, but an identical set of non-moral facts at our world do not determine an identical set of moral facts, then it cannot be the case that the non-moral facts at *W wholly* determine the moral facts at *W*, contravening *Supervenience.* But again, this argument will be question-begging if it simply asserts *Strong* without first dealing with the error theorist's counterargument.

A modal argument against error theory that was not question-begging would need to start by showing that error theory does not entail that moral facts are impossible. In fact, this is what Christian Coons has attempted to do in a recent paper (Coons 2010). With one exception, Coons claims that the arguments that error theorists have offered in support of their thesis do nothing to call into question the *possibility* of moral facts.[6] He quotes Jonas Olson's summary of Mackie's error theory:

> It is fairly obvious that the argument against categorical reasons that proceeds via Mackie's queerness worry and debunking explanations of moral beliefs is based on an appeal to Occam's Razor. The gist of the argument, after all, is that error theory offers a theoretically simpler and hence preferable explanation of the phenomenon to be explained (i.e. moral thought and talk) than do competing realist explanations. (Olson 2010: 67)

From considerations like this Coons concludes that error theorists do not have a case against the possibility of moral facts:

> For if error-theorists felt they already had reason to indict the possibility of moral facts, then appeals to debunking explanations of moral beliefs and theoretical simplicity would be superfluous. (Coons 2010: 7)

And from the possibility of moral facts and *Supervenience* he concludes that there are moral facts at our world and thus error theory is false.[7]

In making his claim that moral facts are at least possible, Coons explicitly assumes both that (a) "moral claims are coherent and … we cannot rule out a priori that they can be true" (Coons 2010: 7); and that (b) "no non-moral features of the actual world preclude the instantiation of moral properties" (ibid.: 5). Of course, it would be an endless task to defend

---

[5] The nature of the supervenience relation between moral and non-moral properties is controversial, particularly with regard to what exactly is supposed to supervene on what. This is not surprising given the broad acceptance of *Supervenience* coupled with the diversity of accounts of what our moral judgements commit us to, ontologically speaking. Nicholas Sturgeon (2009) has argued that given the different ontological commitments of rival metaethical theses, there is in fact no one supervenience thesis that is generally agreed upon. I have stated *Supervenience* as blandly as possible to avoid such subtleties.

[6] The exception is Richard Joyce.

[7] His actual argument is more complex than this, because he uses a modal argument not only to attack error theory, but also to demonstrate the truth of a substantive moral claim (roughly, that causing harm in a way that advances no one's ends is always wrong). But his more complex argument contains the simpler argument that I discuss here, so the former fails if the latter does.

all the assumptions one makes in formulating a philosophical argument; but these seem to be very big assumptions to make in the context of *this* discussion. Note that (a) is tantamount to *Weak.* If (a) is false, then error theory is true; and we have seen in the previous section that both Streumer's and Loeb's versions of the error theory entail that (a) is false (neither of which, it should be noted, appeal to debunking arguments or theoretical simplicity). So anyone tempted by a modal argument against versions of the error theory that entail (a) is false, must show why such theories fail to establish the falsity of (a). To assume that (a) is true is question-begging in this context.

Indeed, we can catch Coons at this. Coons singles out Joyce as the one error theorist whose arguments seem to entail that moral facts are impossible. I don't think he's entirely right about this, as I argued in the previous section that Joyce's error theory does not rule out *Weak.* But leaving that aside, why doesn't Coons see the need to engage with Joyce's argument? This is part of what he says in response to the challenge from Joyce:

> In order to mount his case against the possibility of moral requirements, Joyce must assume that my arguments cannot work. For if we can *demonstrate* that certain actions are morally required, then there *will* be convergence in agents‟ practical reasons. Each person susceptible to the demonstration has a practical reason to abide by the requirement. After all, one's susceptibility reveals commitment to the moral requirement, even if one does not antecedently accept or believe it. (ibid.: 12-13, emphasis original)

Coons says that Joyce must assume that his (Coons's) arguments cannot work. Joyce should reply that it is Coons who must assume that his (Joyce's) arguments cannot work! Coons does nothing of substance to demonstrate that certain acts are morally required because his argument for this depends entirely on the unsupported assumption that such requirements are possible, an assumption which has been positively argued against by Joyce (at least with respect to this world). Thus if Coons wishes to refute Joyce he has to deal with the substance of his arguments.

In the context of this discussion, (b) also seems a gratuitous assumption. Coons uses (b) to support his premise that "There is a possible world, like ours, where error theory is false", which is true iff *Strong* is true. But (b) is really just equivalent to *Strong*. Saying that nothing precludes the instantiation of moral facts at our world amounts to the same as saying that moral facts are possible at worlds non-morally like ours. The first premise of Coons‟s argument is again just an unsupported assumption. This wouldn't be so bad if it was one that error theorists were also compelled to make. But as I've shown, not only are they not compelled to make this assumption, their arguments, if sound, entail precisely the opposite.

We can now see that Coons's claim that appeals to debunking arguments and theoretical simplicity would be superfluous if error theorists had the resources to indict the possibility of moral facts is misguided. It seems that Coons makes this claim because he fails to notice that these arguments can be used to support the claim that moral facts are impossible (either at all worlds, or at least all worlds non-morally identical to our own). When discussing the argument from relativity in the previous section, I argued that if the best explanation for the diversity of moral opinions at our world was the one that Mackie gave, then this would also be the best explanation for the diversity of moral opinion at any non-moral duplicate of our world; after all, the only factors that we can consider in these situations, if we are not to beg the question against error theory, are the non-moral ones, which are identical in each case. And when discussing the argument from queerness, I argued that if it was most reasonable to accept a debunking explanation of moral discourse at this world, then it would be most reasonable to accept a debunking explanation of moral discourse at a non-moral

duplicate of this world. These debunking arguments (which contain within them appeals to simplicity) are not superfluous if we can use them to establish the impossibility of moral facts.

### 3. A final argument from Coons

Before closing, I would just like to comment on another argument against error theory that appears at the end of Coons's paper, which like the argument we've been discussing up till this point attempts to derive a substantive moral conclusion from what Coons takes to be uncontroversial claims, acceptable to the realist and error theorist alike (Coons 2010: 13-15). The argument goes:

> (i) If you are faced with two options, *A* and *B*, and *A* is morally forbidden if anything is, and *B* is not morally forbidden if anything is, then you morally ought not to do *A*.
>
> (ii) If anything is morally forbidden, then accelerating a car into a crowd of people for no good reason would be, and refraining from doing so would not be.
>
> (iii) Thus accelerating into a crowd for no good reason is morally forbidden.

This argument can easily be shown to be flawed. I'll grant Coons (ii), though I have reservations about it, but (i) is simply not true. Coons says that (i) is a corollary of the claim that "one morally ought to avoid doing what is morally forbidden", but it clearly isn't. "One morally ought to avoid doing what is morally forbidden" is uncontroversial because to say that something is morally forbidden is just another way of saying that you morally ought not to do it. Thus (i) could be reworded as:

> (i') If you are faced with two options, *A* and *B*, and *A* is morally forbidden if anything is, and *B* is not morally forbidden if anything is, then *A* is morally forbidden.

If we remove the distracting 'option *B*', which isn't doing any work in this argument, we get the very dubious:

> (i'') If *A* is morally forbidden if anything is, then *A* is morally forbidden.

Clearly Coons intends the antecedent of (i) to strictly imply its consequent, but it is clear enough that the consequent of the improved (i'') does not follow from its antecedent. It is akin to:

> (α) If the keys are in my kitchen if they are anywhere in my house, then the keys are in my house.

The only thing close to (i) that the uncontroversial claim that you ought to avoid doing what is morally forbidden allows us to construct is the unedifying:

> (i''') If *A* is morally forbidden, then you morally ought not to do *A*,

which is equivalent to

(i"") If *A* is morally forbidden, then *A* is morally forbidden.

Error theorists will be quite happy to accept the truth of (i""), but clearly the conjunction of (i"") with (ii) will not give us (iii).

## 4. Conclusion

What should error theorists say about the possibility of moral facts? It depends on which version of the error theory they hold, but – as I showed in section one – several versions of error theory can be seen to imply that moral facts are not possible at a world that is nonmorally identical to our own. That means that any modal argument against any of these error theories will be question-begging.

The positive result of this discussion is that error theories are more robust than they have been taken to be. The claim that no world non-morally identical to ours contains moral facts is stronger than the claim that they do not actually exist, which is all error theorist have hitherto claimed. At the same time, the arguments available to the opponents of error theory remain the same, as modal arguments are not available against it. So error theorists can now make a stronger claim, without leaving themselves any more vulnerable to their opponents.[8]

## References


Ayer, A. J. (1946). *Language, Truth, and Logic* (2nd ed.). London: Gollancz.

Coons, C. (2010) (online publication date). How to prove that some acts are morally wrong (without using substantive moral premises). *Philosophical Studies*.

Cuneo, T. (2007). *The Normative Web.* Oxford: Oxford University Press.

Dreier, J. (2004). Meta-Ethics and the Problem of Creeping Minimalism. *Philosophical Perspectives 18:* 23-44.

Joyce, R. (2001). *The Myth of Morality.* Cambridge: Cambridge University Press.

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

Loeb, D. (2008). Moral Incoherentism: How to Pull a Metaphysical Rabbit out of a Semantic Hat. In W. Sinnot-Armstrong (ed.), *Moral Psychology,* Vol. 2. Cambridge MA: MIT Press.

Mackie, J.L. (1977). *Ethics: Inventing Right and Wrong*. London: Pelican Books.

Olson, J. (2010). In Defence of Moral Error Theory. In M. Brady (ed.), *New Waves in Metaethics*. Basingstoke: Palgrave MacMillan.



[8] I would like to thank David Liggins and Michael Scott for much helpful discussion. This paper was presented at the Stockholm Philosophy Graduate Conference 2011, and I am grateful to Jonas Olson for his comments on that occasion.

Streumer, B. (2008). Are There Irreducibly Normative Properties? *Australasian Journal of Philosophy, 86*: 537-61.

—— (2010) (online publication date). Are normative properties descriptive properties? *Philosophical Studies*.

Sturgeon, N. (2009). Doubts About the Supervenience of the Evaluative. In R. Shafer-Landau (ed.), *Oxford Studies in Metaethics* 4. Oxford: Oxford University Press.

Williams, B. (1981). Internal and External Reasons. In his *Moral Luck.* Cambridge: Cambridge University Press.