SUPERVENIENCE AND MORAL REALISM

ALISON HILLS
University of Oxford

1 INTRODUCTION

According to objective moral realism, there exist moral properties that are not dependent on our beliefs or our attitudes about them. In Mackie’s famous first chapter of his book *Ethics*, he gives three arguments against this theory. His first argument is that the nature and quantity of moral disagreement gives us reason to deny that there are any objective moral properties. His second argument is that the very nature of putative objective moral properties or moral facts – as essentially motivational (or perhaps, essentially normative) – is reason to doubt that any such thing could exist. Moral realists have not been too impressed by these arguments, however. It is perfectly clear that moral properties can coexist with widespread and persistent disagreement about their nature. In response to the queerness argument, moral realists have either opted for naturalism, trying to show that moral properties are not really peculiar at all – in fact they simply are natural properties – or they have conceded that moral properties are non-natural and that at least to that extent they are queer, but have argued that this is not a sufficient reason to deny that they exist.

The third argument, concerning supervenience, has been comparatively neglected, however, though it was taken up by Simon Blackburn, with responses from a variety of moral realists. In the following sections I will make the supervenience argument against moral realism more precise, showing that it must be directed at non-naturalist forms of moral realism. I will consider one possible response to the argument: I will look at the possibility of denying that supervenience is true.
2 THE SUPERVENIENCE ARGUMENT

It is widely, perhaps even universally believed that moral properties necessarily depend on other properties. To be more specific (but still quite rough):

Supervenience: there can be no difference in moral properties without a difference in the natural properties.¹

Supervenience in one form or another is accepted by a range of philosophers who otherwise have quite different methodical views, including G.E. Moore, R.M. Hare, Blackburn, Shafer-Landau, and Ridge. Some of these philosophers accept moral realism, others do not. But according to Mackie’s supervenience argument, moral realists cannot comfortably accommodate Supervenience. The problem is not that Supervenience is actually inconsistent with moral realism, rather that Supervenience generates an explanatory demand which moral realism cannot meet but certain versions of non-realism can. If the argument succeeds, moral realism is at a considerable disadvantage to a theory like Blackburn’s non-cognitivism.

We can distinguish between different versions of Supervenience. I will assume that the necessity at stake is metaphysical:²

¹ What are the natural properties? It is well-known that “the natural” is difficult to define. Typically philosophers understand the “natural” properties as something like: those properties that are studied by science (or perhaps those that would be studied in a future, complete science). Do mental properties – being happy, being in pain, having a desire or intention, deciding to act – count amongst the natural? One might say no, they do not, and deny Supervenience on that basis. Do supernatural properties – being prescribed or proscribed by God – count as natural properties? Again, one might say no, and deny Supervenience for that reason. But I will assume for the purposes of this paper that the “natural” can be defined sufficiently clearly and sufficiently broadly that supervenience is a plausible claim.

² Alternatively, Supervenience could be primarily a semantic thesis about the meaning of moral terms (or the requirements on a competent user of moral terms), with reference to analytic necessity. In Blackburn’s presentation of the supervenience argument, he seems to be discussing analytic necessity, and requiring a moral realist to explain why it is part of our concept of morality that there can be no moral changes without accompanying changes in natural properties. However, it seems to me more difficult for the moral realist to explain the metaphysical connection between the moral and the natural (at least, this is a difficulty for a non-naturalist moral realist), so I will focus on the metaphysical supervenience claim.
Metaphysical Supervenience: it is metaphysically impossible for there to be a difference in moral properties without a difference in natural properties.

But we can distinguish a stronger and a weaker claim, depending on whether we are talking about differences within one world, or between worlds:

Strong Metaphysical Supervenience: it is metaphysically impossible for there to be a difference in moral properties (between any two possible worlds) without a difference in natural properties.

Weak Metaphysical Supervenience: it is metaphysically impossible for there to be a difference in moral properties (within one possible world) without a difference in natural properties.

That there is a dependence between moral and natural properties is not itself inconsistent with moral realism. But a necessary connection between two kinds of property is the sort of thing that calls for explanation and it is cost to a theory if it is committed to such a connection without being able to explain it. The problem for moral realists is that they have not offered an adequate explanation – in fact in many cases they have offered no explanation at all – of why this dependence should hold. There seems to be a particular difficulty in explaining why weak supervenience is true, if strong supervenience is not. Suppose that in World 1, natural property N1 underlies moral property M, so that there is never a change in property M without a change in N1 in that world. And suppose that in world 2, natural property N2 underlies moral property M, so that in world 1, there is never a change in property M without a change in N2. According to weak supervenience, there can be no “mixed worlds” in which a change in M is sometimes accompanied by a change in N1 and sometimes by N2. But why not? Moral realists have some explaining to do.

3 Blackburn thinks that his non-cognitivism can explain this better, because on that view the function of moral discourse is to praise, recommend or condemn actions, states of affairs etc in terms of their natural properties, in order to regulate our behaviour. So it is clear why weak supervenience is true, given non-cognitivism. But since there is no single best way of carrying out this function, strong supervenience is not true. There is nothing mysterious to explain.
According to some moral realists, moral properties simply are natural properties. So it is obviously true that necessarily, there can be no change in moral properties without a change in natural properties – a change in moral properties is a change in natural properties. In which case there is no need to offer a special explanation for the supervenience of the moral on the natural.4

But non-naturalist moral realism is in a more difficult position. According to non-naturalists, moral properties are not natural properties: the two are distinct. So why is there a necessary connection (either strong or weak) between the two? Non-naturalists need not offer a naturalist explanation for this connection (they need not commit themselves to what has been called “superduper supervenience”).5 But it remains true that they think that there is a metaphysically necessary connection between two distinct kinds of properties, natural and moral, and it is not unreasonable to expect this sort of connection to be explicable.

The non-naturalist might simply bite the bullet and refuse to give any explanation at all. It is simply a brute fact, she might say, that there is this necessary connection between moral and natural properties. This rejection of the demand for explanation is not very appealing, but the non-naturalist might bolster her position by claiming that others are in an equally difficult place. For example, she might cite other topics where realists have accepted supervenience theses, including the thesis that mental properties de-

---

4 Naturalist moral realists can claim that there are analytic connections between moral and natural predicates, but more typically they deny this, claiming that moral terms cannot be analysed in natural terms, but nevertheless they refer to natural properties (on an analogy with the identity between water and H2O – water cannot be analysed as H2O, but nevertheless water necessarily refers to H2O). Analytic naturalism is not very plausible, since it has not proved possible to provide naturalistic analyses of moral terms; the non-analytic version is more appealing. However, there are serious questions over whether the semantics of moral terms really are similar to those of natural kind terms and it is not obvious that either type of naturalism is sustainable.

5 Superduper supervenience is “ontological supervenience that is robustly explainable in a materialistically explainable way.” (Horgan 1993 p.566). This is more than a moral realist needs to accept, unless they are committed to naturalism. But of course, if they accept non-natural moral properties and they accept a necessary connection that is not naturalistically explicable, they violate naturalism twice over. But if naturalism is false, as the non-naturalist believes, two departures from naturalism are no worse than one. So it is not a compelling objection to a non-naturalist moral realist that she has no naturalistically respectable explanation for supervenience.
pend on physical properties, that chemical properties depend on atomic properties, that colour properties depend on physical properties. It is rare for anyone to offer explanations for the proposed necessary connections between these sets of properties. Yet philosophers advocating these supervenience theses are often realists about the supervening properties. So a moral realist who accepts an unexplained dependence between the moral and the natural is no worse off than these.

This defence of moral realism is vulnerable, however, to someone prepared to be sceptical about all these topics – or rather, someone who would be sceptical about an unexplained necessary connection between the mental and the physical, for example. They might insist that we must be either (type) identity theorists, or irrealists about the mental. In which case, an appeal to “companions in guilt” will not save the moral realist.

4 DENYING SUPERVENIENCE

We have seen that the Supervenience Argument fails against naturalist moral realism – for they can explain the dependence of the moral on the natural easily, because moral properties are natural properties. Non-naturalist moral realism (which was Mackie’s original target anyway) does not fare so well.

I want now to consider a radical option for non-natural moral realism in response to the Supervenience Argument. Shafer-Landau describes this response as “extreme” and no one to my knowledge has taken it sufficiently seriously even to explore it. This response is to deny supervenience.6

Supervenience, whether strong or weak, is a modal claim. It states that there could be no moral difference (within one possible world, or at all) without a natural difference. I suggest that a non-naturalist realist might deny this modal claim and make a much weaker claim about the actual world:

Constant conjunction: in the actual world, there are no differences in moral properties without differences in natural properties.

---

6 There are other responses open to a non-naturalist, for example Shafer-Landau (2003, pp.80–115). For criticism of this defence of non-naturalism, see Ridge (2007).
In other words a moral difference is constantly conjoined, in the actual world, with a difference in natural properties. But there is no implication that this is so necessarily, or that it must be so in every possible world.

There is a danger that constant conjunction as stated here is a trivial truth. For after all, it is plausible that no two distinct actions have exactly the same natural properties – one must be at a different time, or in a different place, or carried out by a different agent. But there may be a more interesting class of natural properties – such as those connected with welfare or happiness – such that in the actual world there is no difference in moral properties without a difference in those natural properties. This revised version of constant conjunction, that is not trivially true, I will consider as a rival to Supervenience.

**Constant conjunction (revised):** in the actual world, there are no differences in moral properties without differences in (some interesting subset of the) natural properties.

The benefits to the non-naturalist of denying supervenience are obvious. There is no metaphysical necessity to explain. The Supervenience Argument must fail. Of course, you might say that there is no point in a non-naturalist moral realist denying supervenience when they are already willing to depart from naturalism by advocating “queer” non-natural properties, and more generally, they are prepared to take on the commitments of common sense morality even at the expense of considerable metaphysical concerns. But even if in general you think that common sense morality can trump metaphysics, in the sense that we should accept the existence of “queer” non-natural properties because common sense morality is committed to them, it does not follow that you should shoulder extra metaphysical burdens when you could just as easily refuse them. And it is quite clear that accepting extra necessary connections is an extra metaphysical cost.

If there is only constant conjunction between two distinct sorts of property, however, and no necessary connection at all, why has everyone thought that there is one? We might appeal to Hume, writing in a very familiar way about a similar problem. According to Hume, we see that cause and effect are constantly conjoined, but we do not see any necessary connection between the two. Nevertheless, we think that there is a necessary connection:
“after a repetition of similar instances, the mind is carried by habit, upon the appearance of one event, to expect its usual attendant, and to believe that it will exist.” (Hume, *Enquiries*, p.75).

We have a feeling of determination, of expectation at seeing the effect when we have seen the cause, that is the explanation of our belief that cause and effect are joined by a necessary connection.

Perhaps we can give a similar explanation of our commitment to the supervenience of the moral and the natural. We see that the moral and the natural are constantly conjoined, so that we come to expect to see a difference in the interesting natural properties when we see a moral difference. It is the feeling of expectation that explains our belief that there is a necessary connection between the two. In fact, there may be no necessary connection at all.

But it is not obvious that we can borrow a Humean explanation of our belief in necessary causal connections to explain our commitment to Supervenience. There are important differences between the two. In the first place, we can conceive of causal connections failing to hold. We can imagine that one billiard ball hits the other, and the second explodes, or changes colour, or bursts into flame. But it is not so clear that we can conceive of moral differences without natural differences. And indeed this feature of moral properties has been used in an argument in defence of Supervenience. I will assess this argument in the next section.

5 Conceivability

Though many philosophers seems to accept Supervenience in one form or another, arguments for it are rather thin on the ground, perhaps because many people have found the claim so obvious as to not need one (and because no one has seriously argued against it). Blackburn merely says that the supervenience of the moral on the natural is “widely held” and later that to deny it would be a logical mistake. (Blackburn 1993 pp.114, 116, see also p.137.)

Mike Ridge does offer something more in defence of supervenience (his account of supervenience is called S, which in our terms is a version of strong metaphysical supervenience):

(S) is extremely plausible, to the point that someone who denied it would thereby betray incompetence with normative concepts. To deny (S) would be to al-
low, for all that has been said so far, that it could have been the case that the world was exactly like the actual world in all of its non-normative and descriptive features, yet Hitler’s actions were not wrong. Since all the non-normative and descriptive facts are the same in this possible world it will still be true that Hitler killed the same people, had the same intentions, etc. Such bare normative differences seem *inconceivable*. (p.335).

There are plenty of ways that this argument could be questioned. First, one could deny supervenience without accepting that any natural property could be associated with any moral property. So one might deny supervenience without accepting that there is a possible world in which what Hitler did was morally right. But then one would have to explain why this natural property could not be associated with moral rightness, and we have taken on a different but not obviously easier explanatory burden. So I will not question that premise of the argument.

What I really want to look closely at is the connection between conceivability and possibility, or rather on inconceivability and impossibility. I think that this is at the root of many people’s acceptance of supervenience, so if the non-naturalist can cast doubt on that connection in this context, she will have made good progress towards denying supervenience.

The connection between possibility and conceivability is notoriously problematic. But rather than talk about general issues, I want to focus on the specific link between conceiving of moral properties and impossibility. The first problem is that it is not at all clear what it is to conceive of a possible world with or without certain moral properties, for example, what exactly is it to imagine that Hitler’s actions are or are not morally wrong in some world? If we cannot make good sense of conceiving of moral properties, or failing to conceive of them, the conceivability argument for supervenience obviously fails.

For the purposes of argument, however, I will assume that we can make sense of conceiving that moral properties exist or do not exist in a world. Nevertheless, it is well-known that there are interesting and puzzling issues surrounding conceiving of moral properties. In particular, we seem to find unusual difficulties in imagining that what we believe to be morally wrong is (in some world) morally right. These go under the title

---

7 Ridge’s (S) is his own version of Supervenience: (S) Necessarily: Two entire possible worlds cannot differ in their normative properties without also differing either (a) in their non-normative properties or (b) in their descriptive properties. The differences between this thesis and the ones I have set out are not relevant for our purposes, and so I will not be discussing them.
“puzzles of imaginative resistance” and are typically focussed on a discussion of fiction. Whilst some of the puzzles raise issues specific to fiction (such as questions about the authority of a fictional narrator), some of them are more general, puzzles about imagining worlds that need not be fictional in the sense of being explicitly part of a fictional narrative.

Can you imagine Hitler’s actions (or actions that are naturalistically identical to Hitler’s) being morally right? Probably not. Similarly, you probably cannot imagine that it is morally right to kill a baby because it is a girl, or that genocide and slavery are morally acceptable.

The key question for us is: what is the explanation of this inconceivability? One answer is precisely that we cannot imagine these because they are impossible. Walton explains the issue as follows:

Moral properties depend or supervene on ‘natural’ ones … being evil rests on, for instance, the actions constituting the practices of slavery and genocide … This … accounts … for the resistance to allowing it to be fictional that slavery and genocide are not evil … Our reluctance to allow moral principles we disagree with to be fictional [that is: true in the world of some fiction] is just an instance of a more general point concerning dependence relations of a certain kind

And further:

We need an explanation of why we should resist allowing fictional worlds to differ from the real world with respect to the relevant kind of dependence relations. My best suspicion … is that it has something to do with … an inability to understand fully what it would be like for them to be different (Walton, 1994, p.46).

We are struggling to imagine something that we think is impossible, that what Hitler did was morally right. We simply cannot understand how that could possibly be true.

If this explanation of our imaginative failure is right, we have some evidence that we are committed to a metaphysically necessary connection between the moral and the natural. Whether this commitment is evidence that there is such a connection is a separate question, one that I will not pursue here. Instead I want to consider, drawing on work by Tamar Gendler, whether there is a very different explanation for our imaginative failure.

8 Walton (1994, pp.43–46).
Gendler suggests that, in general, it is possible to imagine that necessarily false propositions are true. But most important is her positive account of the source of the imaginative failure:

Classic imaginative resistance arises when a reader can’t imagine a certain moral claim being true in a story (Imaginative Barriers) because she won’t bring herself to adopt the requisite set of generation principles governing the use of moral appraisals (Imaginative Impropriety). So classic imaginative resistance arises when we can’t because we won’t. (Gendler, 2006, p.164).

Gendler suggests that we think there is something wrong with engaging in an imaginative way with the possibility that murder, genocide and slavery are right, for example. Imagining is different from merely supposing (for the sake of argument), which itself can be somewhat distasteful when what we are asked to suppose is something that we regard as very wrong. Imagining requires more engagement than supposing, however, and is therefore more problematic.

Why don’t we want to imagine that something that we think is morally wrong is in fact right? Gendler does not really give an account, but there are a number of possibilities. Perhaps we think it is somehow contaminating to engage with the possibility that it is acceptable to commit mass murder, for example. Or perhaps we think that thoughts and emotional responses associated with morality are habit-forming, and we should not risk forming bad habits. These reasons do not seem to be very strong, but it is quite possible that the desire not to imagine these things is very powerful, even though the reasons for that desire are weaker.³

I suggest that support for the “desire” account of resistance comes from reflection on different false moral claims. Consider something at issue between two of the major ethical theories, for example,

³ Is the desire not to imagine that what Hitler did was right sufficient to explain our inability to imagine it? Gendler herself is tempted by a complication. She suggests that we may think that moral claims are metaphysically necessary, so if we think a claim is false, it is impossible. Hence there are two sources of resistance: a desire not to imagine that it is true and an inability to imagine in detail how it could be true (she thinks that the desire not to do so may lead us to notice the incoherence of the claim itself). But this doesn’t seem to be an essential part of her explanation of imaginative resistance, and she herself is not sure that it should be a part (at least of all cases of imaginative resistance): “isn’t it because I’m not willing to let myself imagine this that I can’t see a way for it to be true?” (Gendler, 2006, p.171).
The benevolent lie: Mary told Peter that his prognosis was good, even though she knew that it was not, in order to make him happier in his last days.

Suppose that you think (like the utilitarian) that it is not always wrong to tell a benevolent lie of this kind. Can you imagine that it is always wrong (as the Kantian believes), that it is a failure of respect for that person not to tell them the truth? Surely you can. Similarly, if you believe that the Kantian is right about this, you can imagine that she is not.

But if so, then you can imagine two worlds with the same natural facts (someone tells another a proposition that they believe to be false in order to make her happy), one in which it is morally wrong, one in which it is morally acceptable. Perhaps in one world utilitarianism is true, in another Kant’s moral theory is true.

There are two aspects to the benevolent lie example which make it easy to imagine: first, there are well-known moral theories supporting different sides, so it is quite easy to see how this particular claim would fit into a wider theory. Secondly, you may think that it is not always wrong to tell a benevolent lie of this type, but it is not so very far off the mark to think that it is always wrong – certainly not so far off the mark as thinking that what Hitler did was morally acceptable. The more mistaken the moral claim is that we are asked to imagine, the more resistance we feel to doing so.

So there is certainly no problem in supposing for the sake of argument that Kant’s moral theory is true, and very little problem in imagining that it is true. And the same for utilitarianism. It is by contrast very difficult to imagine that Hitler was right and not that easy to suppose for the sake of argument that he was.

These are cases in which you are imagining that something you believe to be morally wrong is really morally right (or acceptable). But you can imagine actions that you believe to be morally acceptable are morally wrong too. Suppose that you think that early abortion is morally acceptable. Can you imagine that it is not? I think so – quite easily. Suppose that you think that eating animals is morally acceptable. Can you imagine that it is not? Again, this seems to be quite easy to do. There is no resistance.\(^\text{10}\)

\(^\text{10}\) There are, I think, some difficulties in imagining that something totally morally neutral – like turning on your computer – is morally wrong, partly because it is hard to see how it would fit in with an overall moral theory. But supervenience is not required to explain this.
If the “impossibility” explanation of imaginative resistance were true, it ought to be as difficult to imagine the benevolent lie being always wrong as to imagine that Hitler’s actions were right – for both are metaphysically impossible (and both violate strong supervenience, if in this world Hitler’s actions were wrong and benevolent lies are sometimes right). Since it is harder to imagine some of these examples than others, the “desire” account should be preferred.

A defender of the “impossibility” interpretation might respond by giving an account of how some things that are metaphysically impossible are easier to imagine than others. I don’t think that it is out of the question that we could give such an account – but even if we could, it would still follow that the bare fact of something’s being metaphysically impossible does not explain our resistance to imagining it. Imaginative resistance is not explained by our commitment to supervenience as such.

In the absence of this sort of explanation, the “unwillingness” account of why we are resistant to imagining different moral properties associated with the same natural properties is more compelling. In which case, it is not legitimate to cite the impossibility of conceiving of this as grounds for its being impossible. In the first place, as in the benevolent lie example, we can conceive of this. Secondly, in cases when we cannot, the best explanation seems to be that we are unwilling to do so, rather than that to do so is impossible. So we have no grounds for thinking that it is impossible. This argument for strong supervenience fails.

So far we have set out a thought experiment against Strong Supervenience – the claim that if N1 underlies M1, it does so in every possible world – not Weak supervenience – the claim that if N1 underlies M1 in one possible world, it always does so in that world. Consider a possible world in which there are two benevolent lies. Can we suppose for the sake of argument that first utilitarianism is true and the first lie is morally acceptable, then that there is a change in distribution of the moral properties so that Kant’s moral theory is true and the second lie is morally wrong? Can we imagine that this is true? I think that, with the background of non-naturalism, this is imaginable. If this is right, then this is reason to think that the conceivability argument does not support weak supervenience either.
I have argued that a non-naturalist moral realist should consider responding to the supervenience argument by denying supervenience, that is, by denying that there is any necessary connection between moral and natural properties. In the last section, I showed that one of the main arguments for supervenience – that it is not conceivable that there might be a change in natural properties without a change in moral properties, therefore it is impossible for there to be such – is not compelling. It does not of course follow that supervenience is false; there might be some strong argument for it that I have not considered (though there are other arguments for supervenience in the literature, I think that the conceivability argument is the strongest that there is). I am not suggesting that supervenience is false because we can imagine a change in moral properties without a change in natural properties – for we may be able to imagine things that are in fact impossible. But given that there is a good explanation of our imaginative failure (when it occurs) that does not appeal to supervenience, the conceivability argument for supervenience fails and we have, as yet, no reason to accept it. Without any such reason, a non-naturalist moral realist can simply that there is a necessary connection between (what she regards as) two distinct properties.

So are there any reasons for a non-naturalist moral realist not to accept the constant conjunction account? There are a number of objections that might be raised to it and I will consider two here. First, you might wonder whether the constant conjunction (in this world) between (certain) natural properties and moral properties needs explanation. If it does need an explanation, then the non-naturalist moral realist may not have been made any better off by accepting the constant conjunction account over supervenience.

The basic constant conjunction claim, that there happens to be no change in moral properties in this world without a change in natural properties is likely to be a trivial truth, for any two actions will have different natural properties (at least in terms of the time they take place, where they occur and the identity of the agent and those affected by the action). Clearly a trivial truth like that requires no explanation. Suppose that we restrict our attention to an interesting subset of the natural properties. Do we need an explanation for why there happens to be no change in moral properties without a change in those natural properties in this world? I do not think that a mere regularity, with no implications that the connection is neces-
sary, does require an explanation. It is simply a brute fact, admittedly one that makes a certain kind of moral reasoning and moral argument possible. But a brute fact that does not require explanation nonetheless.

The second problem is potentially more serious. Is the constant conjunction account consistent with moral epistemology? What is the point of thought experiments for example, if you are not finding out about a necessary connection between properties? How could reflection on moral problems possibly be a reasonable strategy if you are not reflecting on a necessary connection?

It is certainly true that reflection and thought experiments play an important role in moral epistemology. But experience also plays an important role; it is sometimes true that one cannot appreciate whether an action is right or wrong without being the agent or the victim. And when one is acquiring moral beliefs in the first place, testimony from one’s parents and guardians, friends and peers is very important.

According to constant conjunction accounts, when you are reflecting on moral problems or considering thought experiments, you are making use of a regularity that exists in this world between moral and natural properties. This regularity you may first discover through experience or through the testimony of others. Then you can use your knowledge in thought experiments to consider possible situations that you have not met – just as an experienced engineer might be able to use reflection or a thought experiment to tell you whether a bridge would stand up if it were built.

Of course there remain significant questions of moral epistemology for a non-naturalist moral realist, in particular, they need to offer an explanation of how we can ever discern non-naturalist moral properties or facts. But this is plainly a serious difficulty for a non-naturalist moral realist whether she accepts supervenience or constant conjunction. So the constant conjunction account does not raise any significant problems of moral epistemology, over those that the non-naturalist already has.

7. CONCLUSION

A non-naturalist moral realist could accept that (a certain subset of) the natural properties are constantly conjoined with the moral properties, rather than that there is a necessary connection between the two. So there is no argument against non-naturalist moral realism on the basis of supervenience.
REFERENCES


