

# Brutalist Non-naturalism and Hume's Principle

Nick ZANGWILL

## Abstract

Does moral non-naturalism have a problem with supervenience? That is, are necessary relations between moral and natural properties mysterious if those properties are distinct? Here I try to remove anxiety about the modal comments of moral non-naturalism. I also want to understand the source of the anxiety for those afflicted by it. That source is a commitment to what is called 'Hume's Principle'. I attack that principle.

## 1. Non-naturalism and supervenience

In order to orientate the discussion, I begin with some elucidatory comments on these ideas.

We may take 'moral realism' to be the view that some positive moral properties are actually instantiated. So there are some positive moral facts. Such facts are judgement-independent in some suitable sense.

We may take 'non-naturalism' to be the moral realist view that some positive moral properties are actually instantiated, so there are some positive moral facts, which are judgement-independent in some suitable sense, and those properties or facts are not identical with, or reducible to, natural properties or facts.

I shall not worry too much about the definition of 'moral' and 'natural' properties. It will be enough if there are some clear cases of each kind. Whether or not the distinction is clear-cut or blurred, with some cases that are hard to classify, does not matter. If 'rude' and 'brave' are hard to classify, then we can ignore those cases and concentrate on those that are easy to classify. Virtue, justice and evil are clear examples of moral properties, while psychological properties, such as having desires, intentions or emotions, are clear examples of natural properties. This means that the psychological properties of a non-physical being, such as a deity, count as natural properties. Physical properties also count as natural properties. Let us not worry about the rationale for this notion of 'natural' property. (I investigate these issues in Zangwill 2017a.)

'Moral-natural supervenience' denotes a systematic set of necessary connections between moral and natural properties and facts. One way of stating supervenience is to say that if something has a moral property, then it has some natural property such that anything that has that natural property must also have the moral property. Or, in another form: if one thing has a moral property and another thing lacks it, then there must be some natural property that one has and the other lacks.

It is widely thought that non-naturalist moral realism has a problem with supervenience principles of this kind. This I deny, and I try to explain why the worry is misguided.

We should distinguish metaphysical puzzles about supervenience, on the one hand, from conceptual and epistemic puzzles, on the other. One purely metaphysical question is: why are there any necessary connections between moral and natural properties if they are not identical? A related purely metaphysical question is: why are there the particular necessary relations that there are? The second kind of question is conceptual or epistemic: why does supervenience hold with 'conceptual necessity'? That is, why is accepting moral supervenience part of what it is to think in moral terms? Or: why is supervenience conceptually necessary when no particular relation between moral and natural properties is conceptually necessary (Blackburn 1993 [1985])? If so, how do we know particular moral-natural necessities, if not conceptually? There are these two broad kinds of issues, though no doubt, in practice, arguments concerning one may bear on the other. One thing to watch for, however, is slipping from one to another in argument, as though they are the same. In this paper, I will focus exclusively on metaphysical issues. (See Zangwill 1996 on the conceptual puzzle.)

One might want to explain the general supervenience principle that whenever any moral property is instantiated, there is some necessary connection between that moral property and natural properties. Or one might want to explain specific moral-natural necessary connections. I shall focus on the general supervenience principle rather than specific moral-natural connections, for the following reason. Assume that we think that promise-breaking by itself is necessarily wrong. (If you think that this connection is not necessary, because sometimes it is alright to break promises, then substitute another example, such as maximizing happiness, or whatever you think is intrinsically right or wrong.) Suppose we ask: why is promise-breaking necessarily

wrong? That question risks ambiguity. One issue is about why it is that promise-breaking, in particular, is wrong. The other issue is about why, given that it is wrong, it is so necessarily. If the latter is the issue, then the explanation of the necessity (not the mere fact) of the wrongness of promise-breaking is just the obtaining of the general supervenience principle plus the fact that promise-breaking is wrong (cf. Salmon 1981). For this reason, explaining the general supervenience principle is more fundamental than explaining the various particular moral-natural necessity relations that there are. They are all well explained by nonmodal moral facts together with the general supervenience principle.

Our task, then, is to explain the supervenience principle given non-naturalist premises. Before I proceed to this task, there is one idea that I want to put to one side. It is an idea that many have been tempted by, which is that there can be conceptual explanations of necessities. The idea seems to be that a conceptual connection between the concepts of As and Bs is part of an interesting explanation of necessary connections between As and Bs. It is difficult to see how this has any plausibility. Concepts refer, whether they are mental items or abstract objects. There may indeed be necessary connections between the two things or properties that two concepts refer to. But any facts about the concepts, even the fact that those concepts are necessarily connected, are too extraneous to the necessities to be explained to be even a candidate for something that might explain the necessary connections between the objects or properties referred to. Consider a clichéd example. Why is it necessary that bachelors are unmarried men? The answer cannot be a necessary connection between two concepts! That would be a bizarre proposal. Instead, the answer is that one property is conjunctive, and the other is a conjunct of that conjunctive property. Concepts do not come into it. Even if we suppose that the concepts necessarily have their property references – perhaps because of ‘externalism’ about semantic content plus the necessity of origin – it makes no difference. There are, then, necessary connections between the concepts because there are necessary connections between their references. But the conceptual necessary connections do not explain the necessary property connections. The idea of conceptual explanations of necessities was quite popular in the early twentieth century (perhaps out of desperation), and the idea has even been revived recently, which I find surprising. But the idea of a conceptual explanation of necessities that connect non-conceptual realities seems seriously confused. For this reason, I only consider metaphysical arguments and explanations in what follows.

## 2. Brutalism

One response to the metaphysical issue of why the supervenience principle holds is to say that the necessities of that principle are brute – which in a sense is a rejection of the question. Some necessities are explained by others, but some are basic. This was Ian McFetridge’s response (McFetridge 1985). I shall call it ‘brutalism’. Not everything is explained by other things; some things just sit there. Why cannot moral-natural necessities be like that?

In my view, however, more can be said, although there is something deeply right about McFetridge’s bullish response. I suggest that the terminus of explanation is the kind of properties in question: it is of the essence of these kinds of properties to have a certain modal structure (Zangwill 1996, 2005, 2008). In particular, an actual instantiation of a moral-natural (= “M-N”) property combination by a thing at a time, brings with it cross-object, cross-time, and cross-world necessitation relations. This brutality is ‘Property Brutalism’ rather than McFetridge’s ‘Necessity Brutalism’. If, at this point – when we have appealed to the nature of moral properties – the question ‘Why?’ is asked, this is one question too many. Why are the moral properties like that? Why do they have that modal structure? That is an odd question. They just do. This is like asking why the fundamental laws of physics are as they are. Different kinds of properties have different essential properties, and moral supervenience just happens, as it were, to be an essential property of moral properties. Physical or colour properties are not modally structured like moral properties and that is a basic fact. Compare: why are numbers not in space and time? There is not much one can say about such questions. At a certain point metaphysics is descriptive and explanations come to an end. The Property Brutalist says that this is the situation with moral supervenience – this is what moral properties are like. The nature of moral properties explains the pattern of necessities. Thus – argues the Property Brutalist – there is no metaphysical problem with moral supervenience for a non-naturalist moral realist.

It might be replied that Property Brutalism does not deal with what is puzzling about the way that necessities bind M and N properties. Suppose one thing at a time has moral property M and natural property N. (This natural property N is not a maximally specific property, such that only one thing could have it, but it is complete enough to be the basis of the M instantiation.) Now suppose also that there is some other thing, or

the same thing at another time, or a merely possible thing that is just like the first thing in natural respects (it is also N). Then, assuming no defeating properties, it too must be M. Or suppose that one thing is M while another thing is not-M. Then there must be some natural property N that one has and the other lacks. But, it might be asked: how can it be that how it is with one thing at a time exerts a power over other distinct things, or over that thing at distinct times or over merely possible things? This seems a strange kind of action at a distance. This is not just a distance between two distinct properties, but a distance between the instantiation of M and N properties by distinct things, or one thing at distinct times, or an actual thing and a merely possible thing. Is that not still a mystery? The metaphysical puzzle is: why is it that if one thing is M and N, then another thing, perhaps at a different time that is N, or a merely possible thing that is N, must also be M? Or: if one thing is M and another is not, there must be some natural property N such that one thing is N while the other is not N. How do the distinct things, as it were, communicate with each other in this way?

The Property Brutalist will reply to the reply that we should expect this because it is a consequence of the generality of any M-N necessitation relation. Since we are dealing in property-to-property necessitations, they apply to any objects that instantiate the M and N properties. It is not as though there are many distinct M-N relations that get instantiated or not. There is one necessitation relation taking the form: necessarily for all x, if Nx then Mx. That explains why Ma is necessitated by Na, Mb is necessitated by Nb, and Mc is necessitated by Nc, and so on. There are M-N property necessitation relations under which many particular things fall. And it is the necessity of the M-N property relations themselves that are brute. So, what is the problem?

The Property Brutalist non-naturalist does not deny that supervenience can be explained. It is just that the explanation of the necessities of supervenience lies in the non-natural properties in question, in their essences. Explanations of necessities do not always take this form. Consider the necessity that if Socrates exists then  $2+2=4$ . The explanation of that necessity has nothing to do with Socrates. By contrast, the explanation of the necessity that water is H<sub>2</sub>O does lie in the essence or nature of the property of being water. The explanation of moral supervenience is of the latter sort, despite the lack of reduction. The nature of moral properties explains the pattern of necessities. It is the essential nature of moral properties to generate necessities tying moral and natural property instantiations. But if we ask why a property has the nature it does, then one asks a question with no answer. Why cannot numbers have colours and why cannot colours feel pain? A brutalist response to these questions is also plausible.

### 3. Hume's Principle(s)

One way to try to reinstate or resuscitate a problem, in the face of Property Brutalism, would be to insist that there is something mysterious about necessary connections between distinct things, and if the non-naturalist says that action at a distance is just built into the essence or nature of non-natural moral properties, that does not make it any less mysterious. This is just a reason not to accept properties with such essences into our ontology. This leads us to a specific way of trying to capture the alleged metaphysical problem with moral supervenience. Hume famously said that there are no necessary connections between distinct things (Hume 2002, I.III.IV). This has come to be called 'Hume's Principle'. The thought is: if two things are distinct, then it seems that they should be able to vary independently of each other. (See also Wittgenstein 1922: 1.21, who says something similar about facts.) Moral supervenience, for non-naturalist moral realism, in particular, seems a mysterious action at a distance because it violates Hume's Principle. This is Tristram McPherson's diagnosis (McPherson 2012). Perhaps it is for this reason – the violation of Hume's Principle – that moral supervenience, like moral properties themselves, has been thought to be mysterious, strange, weird, peculiar, queer, etc. (Mackie 1977, chapter 1). And perhaps it is this violation of Hume's Principle by moral supervenience that makes the idea of non-natural moral properties particularly mysterious. (McPherson weakens the principle so that it says that postulating brute necessities "counts against a view"; but that does not make much difference.)

McPherson considers various possible ways in which a non-naturalist might try to explain necessary ties between distinct things or properties (McPherson 2012, section 4), and he then argues against those suggestions. But this way of arguing assumes that non-naturalism will be on the defensive and will try to reply to the objection by supplying explanations that respect Hume's Principle.

Instead, Brutalist Non-naturalists should not be embarrassed about brute necessities between distinct moral and natural properties in the first place, and they should not be worried about violating Hume's Principle. So they should not see the need for the explanations that McPherson kindly offers, and then criticizes. The whole

dialectic presupposes that there is something to be said for Hume's Principle. But that should be the controversial point at issue. The dialectic that McPherson assumes should be challenged by Brutalist Nonnaturalists.

The fact is that Hume's Principle is false – clearly false. Let us distinguish strong and weak versions of Hume's Principle. The strong version is cast in terms of some X and Y not being identical while the weak version is cast in terms of some X and Y not being wholly distinct. Both are false, but the strong version is even more obviously false than the weak version. Moreover, Hume's Principle can also be cast in the fact or property mode or in the object mode. In each mode, weak and strong versions of the principle may be distinguished.

In the fact or property instantiation mode, and in the strong version, it says:

It is not the case that there are facts or property instantiations p and q, where  $p \neq q$ , such that necessarily if p then q.

Or in the object mode the strong principle says:

It is not the case that there are objects x and y, where  $x \neq y$ , such that necessarily if x exists then y exists.

In the weak version, the fact or property version says:

It is not the case that there are facts or property instantiations p and q, where p is wholly distinct from q, such that necessarily if p then q.

Or in the object mode, the weak principle says:

It is not the case that there are objects x and y, where x is wholly distinct from y, such that necessarily if x exists then y exists.

As we will see, we need a better grasp of what "wholly distinct" means. But let us first focus on the strong principle.

#### **4. Counterexamples to Hume's Principles: Abstract entities**

Counterexamples? Where to start? It is like being a child in a sweet shop!

(1) Consider sets and their members. These are non-identical. And they are wholly distinct existences. They are even in radically different metaphysical categories in many cases (although sets themselves may also be members of sets). Sets and their members are wholly distinct, not merely not identical. But sets and their members are necessarily connected, contrary to Hume's Principle(s). The existence of the things, the members, necessitates the existence of the set of those things. If so, why not natural facts and non-natural moral facts? The combination of necessarily connected wholly distinct existences is uncontroversial for sets and their members, so why not also for moral and natural property instantiations?

(2) Next, consider numbers, as conceived by Platonists about Mathematics. Different numbers are non-identical. And on most views, they are also wholly distinct. Nevertheless, on most Platonist views, there are necessary connections between the mathematical entities, even though the entities are distinct. For example, it is necessary that 4 lies between 3 and 5. And 4 is necessarily non-identical with the number 3 (and wholly distinct from the number 3). Nevertheless, numbers form an infinite sequence, and each member in the sequence is necessarily connected to all the others. Is that strange? Who knows?! Perhaps asking whether it is strange is strange. The point is that the existence of necessary connections between distinct numbers, for most Platonists, is just what it is to be a number.

Suppose a Platonist rejects necessary connections between numbers. That would generate the bizarre possibility that 4 might have been between 7 and 9 rather than between 3 and 5. Could one be a mathematical realist only about even numbers, and think that odd numbers are completely different matter, a mere fiction

of the mind? Or perhaps only the number seven is real and all the rest are illusory? But, for most Platonists, if one exists then they all do. Whether any numbers exist is one thing; their necessary relations to each other are another. If it is said that infinite sequences of necessarily connected distinct objects are mysterious, that looks like a sheer prejudice against the kind of things that numbers are. Again, numbers are wholly distinct existences yet necessarily connected, and that is part of what it is to be a number. (A mathematical structuralist might say that numbers are not merely necessarily related but also essentially related. That makes no difference. Wholly distinct things are still necessarily related whether or not they are also essentially related.)

Jerry Fodor and Ernie Lepore note that if one number exists then they all do, but they draw the wrong conclusion that “being a number is really a relational property” (Fodor and Lepore 1992, 3), whereas the right conclusion is that numbers have their relational properties necessarily. If we want to put it in terms of the property of being a certain number, then we should say that that property is not a relational one; it is an intrinsic property that explains the necessary relational properties of numbers.

Perhaps there could be a view according to which mathematical entities are non-identical but not wholly distinct. On such a view, the relations between them might not only be necessary but also essential, because distinct numbers stand in part-whole relations to each other. Thus 3 and 5 would be parts of 8. On this view, numbers are not wholly distinct from each other, and thus they conform to the weak but not the strong form of Hume’s Principle. However, this is not a standard Platonist view.

These two examples – sets and numbers – are damning for both strong and weak versions of Hume’s Principle and for arguments that assume them.

## **5. Abstract/non-abstract and Weak Hume’s Principle**

We can imagine those who want to argue against non-naturalism from moral supervenience complaining that sets and numbers are special cases – abstract objects – and they might say that special rules apply there. If that is said, we may complain, in reply to the complainers, that moral and natural properties are special too. Those who object to non-naturalism, such as McPherson, have no choice but to concede many violations of Hume’s Principle. Many cases seem to be special. There seems to be a lot of specialness around. So why not tolerate a little more specialness? Why tolerate specialness in so many cases but somehow just not moral-natural specialness. That seems unfair discrimination.

However, suppose that the complainer persisted, saying that the oddness of necessary connections between distinct things is somehow different and worse for spatio-temporal things. This must be the reply, or else the objection to nonnaturalism from Hume’s Principle is completely dead. Deceased. Expired and gone to meet its maker. Pushing up daisies. Kicked the bucket. ... The complaint would have to be that it is especially difficult to see how properties of two distinct things of a non-abstract sort – things with spatio-temporal locations – can be necessarily connected, even if there are such necessary connections between wholly distinct abstract objects, which, it might be admitted, are weird. The thought would be that weird abstract objects may be bound by weird necessary connections between wholly distinct things, but ordinary spatio-temporal things are not weird, at least in that way, and so there is no reason to expect them to obey weird modal principles. Hence, the argument would be that it is particularly violations of Hume’s Principle among non-abstract objects that is disturbing. It might be conceded that Hume’s Principle is false quite generally, although it is true for non-abstract objects.

Discriminating between abstract and non-abstract in this way seems arbitrary. But let us put that to one side for the sake of argument. Even so, this response would only be available if Hume’s Principle has some *prima facie* plausibility for non-abstract objects. Does it? McPherson’s paper assumes that there is some *prima facie* plausibility to the principle; but there is certainly none for abstract objects. But perhaps there is some plausibility for non-abstract objects.

Here again there is a wealth of counterexamples, especially to the strong version of Hume’s Principle. We are back in the sweet shop! Where to start?

But let us first ask what “wholly distinct” means if it is not just non-identity. Many writers make a distinction here, but they seem to use the phrase in different ways.

As a first hypothesis let us assume that “wholly” contrasts with “partly”: things are not wholly distinct when they are related as parts and wholes. London and North London are not identical but neither are they wholly distinct in this sense. But North London is part of London and necessarily so. Thus, this is a non-abstract counterexample to Strong Hume’s Principle, but not Weak Hume’s Principle, which is designed to accommodate mereological relations and facts. (By “part” I mean “proper part”.)

## 6. Non-abstract counterexamples to Hume’s Principles

(3) One source of non-abstract counterexamples to both Strong and Weak Hume’s Principle are cases of constitution. A statue and the clay that constitutes it are not identical, but they are necessarily connected. Perhaps a particular piece of clay is not necessary for the continued existence of a statue. Nevertheless, some physical matter is necessary. So, this is a counterexample to strong Hume’s Principle. Suppose someone says that the statue and the clay are non-identical but they are not wholly distinct. Well, it is now difficult to know how to understand what this “wholly distinct” is supposed to mean. We have some grasp if “wholly distinct” implies a contrast with “partly distinct”. If so, Weak Hume’s Principle covers only things that stand in the identity or part-whole relations. Parts and wholes are non-identical and necessarily connected, because they overlap and share parts. The trouble is that cases of constitution are not related as parts and wholes. The clay is not part of the statue or vice versa, and the wood is not part of the table or vice versa. If so, constitution cases are counterexamples to both Strong and Weak Hume’s Principle, since statues and tables must be physically embodied. Or is Weak Hume’s Principle somehow to become elasticated so as to include not just identity and part-whole relations but also cases of constitution?

(4) Another source of non-abstract counterexamples to Strong Hume’s Principle are determinables and determinates, like scarlet and red, monkeys and animals. These are not identical yet they are necessarily related. Necessarily what is scarlet is red; and necessarily monkeys are animals. Could Weak Hume’s Principle deem these cases somehow not wholly distinct? Well, either “wholly distinct” is understood mereologically, as having no parts in common, or not. If it is, then determinates and determinables are counterexamples to Weak Hume’s Principle, since determinates and determinables are not related as parts and wholes. Again, we need to know more about what “wholly distinct” means. How do we go on applying it to new cases if this phrase is not interpreted mereologically? Both constitution and determinate/determinable relations are counterexamples to Weak Hume’s Principle, unless “wholly distinct” is to be understood in some other way. Or, again, does Weak Hume’s Principle undergo random elastic expansion to include these cases?

(5) Another case is colour exclusion: being entirely red necessarily excludes being entirely green. But these are wholly distinct properties that are necessarily connected in that they necessarily exclude each other. (The case is not unlike the number case.) This was one of the major reasons that Wittgenstein moved from his early Humean Tractarian views. (See, for example, Wittgenstein 1979, 73–77, 149.) A related source of non-abstract counterexamples to Weak Hume’s Principle are non-abstract negative properties that are determined by non-abstract positive properties: for example, the properties of being water and not being copper are not identical yet they are necessarily connected: one necessitates the other. (I have argued elsewhere that negative properties, or logical constructions of negative properties, cannot by themselves, determine positive properties; see Zangwill 2011.) So Strong Hume’s Principle is refuted again. How about Weak Hume’s Principle? Are the properties wholly distinct? It is not obvious. Does being water overlap with not-being- copper? Is not being water part of being copper? That does not sound right at all. It is true that if we draw a Venn diagram, the things that are water and the things that are non-copper do overlap, and the same with red and non-green. The corresponding sets overlap. Yet properties and sets should not be confused, even if some properties determine sets. What does it even mean for two properties to be mereologically related? Consider the property of being such that  $2+2=4$ , which is true of everything. But that does not mean that being blue is not wholly (mereologically) distinct from the property of being such that  $2+2=4$ . The mathematical and colour properties have nothing to do with each other. Similarly, being red and not being green are wholly (mereologically) distinct, even though the sets of objects in the ‘extension’ of each of the properties (the things that possess them) overlap. The Venn diagram objection confuses objects and properties.

(6) Other possible kinds of cases, in principle, would be where two things have a common determination. Perhaps God both moves our minds and our bodies. Mind and body are then necessarily related but wholly distinct existences. Another example of the same sort, which makes assumptions that I cannot defend here, is that of knowledge and truth (see Zangwill 2013). Knowledge and truth are necessarily connected but wholly distinct existences. The combination of beliefs plus facts are common determinants of both knowledge and truth. Beliefs plus facts plus something else (such as the right belief-fact connection) determine knowledge. And beliefs plus facts determine truth. Yet knowledge and truth are completely distinct existences, necessarily related.

(7) Another source of counterexamples is that of properties related in special and basic sciences. Assuming that property reduction implies property identity, then necessities without reduction are common in the (non-abstract) special sciences, as many, such as Jerry Fodor, have emphasized. Fodor pointed to the existence of laws in special sciences without property reduction (Fodor 1981, 1987). Meteorological and geological properties and laws are two of Fodor's examples. Meteorological and geological properties are causally efficacious; and yet they resist reduction to microphysical properties where 'reduction' implies that there are identity relations between these special science properties and microphysical properties. Moreover, special science properties or facts do not stand in a part-whole relation to the microphysical properties or facts. Nevertheless, these properties are usually thought to supervene on microphysical properties. Their causal efficacy is not entirely explained by supervenience, but supervenience seems to be necessary for causal efficacy. Without supervenience, these properties could not be causally efficacious, given other relatively uncontroversial considerations. So, there are necessary connections between wholly distinct properties. That may all be wrong. But it is a widely accepted scenario in the special sciences, which collides with Hume's Principle(s). The anti-brutalist might reply: even if this combination is common, how can it be? Is not the supervenience plus irreducibility combination puzzling anywhere in the special sciences? But this combination is only puzzling once one accepts Hume's Principle(s). Why not keep the special sciences and get rid of Hume's Principle(s)?

(8) There seem to be many other examples: people's actions and their reasons for those actions seem to be distinct but necessarily connected. Or, again: people have necessary origins (see deRosset 2009). So do biological phenomena such as organisms, organism-parts and species. There are also non-biological examples. Granite necessarily comes from volcanic rock. Mountain tarns are necessarily formed by glacial action. Yet the things related are not just non-identical but wholly distinct. The volcano or glacier, millions of years ago, is one thing, the present granite or tarn today is something wholly distinct.

And so on and so on and so on! Common sense and science contains many, many examples of necessary connections between non-abstract things that go beyond identity and non-overlap.

Thus, whether we look at abstract objects or non-abstract objects, Hume's Principles, both Strong and Weak, face a counterexample tsunami! The fact is that Hume's Principle looks increasingly eccentric, unmotivated and implausible.

## 7. Other writers

Before turning to give a diagnosis and a principled argument against Hume's Principle(s), as opposed to appealing to the tsunami of counterexamples, I want to comment on some writers who have participated in recent discussion of moral non-naturalism and Hume's Principle.

(A) In a brief discussion, David Enoch rejects Hume's Principle as a "metaphysical dogma" and he cites Kit Fine's counterexample of sets and their members (Enoch 2011: 147–148). Of course, I think this is right. But it would be best to justify the rejection of Hume's Principle on principled grounds, not just by a counterexample. (Philosophers are perversely disposed to dispute counterexamples, which is why I assembled so many of them.) And perhaps some "metaphysical dogmas" are correct. As we will see below, I think that we can give a principled and intuitive case for rejecting Hume's Principle(s).

(B) Erik Wielenberg argues that Hume's Principle is self-refuting, or rather that asserting it is, because it claims that flouting Hume's Principle supplies justification for rejecting theories. But, argues

Wielenberg, this claim itself seems to flout Hume's Principle because it asserts a necessary relation between the property of flouting Hume's Principle and the property of justifying rejection; but these properties are non-identical (Wielenberg 2014: 33–34). I sympathize with the direction of this clever argument, but the objection is surely too good to be true. For one thing, justification might or might not be a non-natural normative property. Perhaps it is a property that is reducible to natural properties, or perhaps it is not a property at all, as epistemic expressivists like Hartry Field think (Field 2009). If so, asserting Hume's Principle would not be self-defeating. Secondly, why take support relations to be necessary links to the grounds? Indeed, I think it unlikely that they are (Zangwill 2017b). I agree that arguing that we should not think in epistemic terms is indeed self-refuting (Zangwill 2010, sections VI and VII; Zangwill 2017a, section 1.6). But that is a different kind of self-refutation argument. I would agree with Wielenberg that Hume's Principle is self-refuting given the assumptions that epistemic properties are non-natural normative properties, and also that a supervenience principle holds for such properties, and also that support relations are necessary. But if not all these assumption hold, then we have not yet been given a reason to deny Hume's Principle. The very question that is at issue is whether epistemic or other normative properties are non-natural properties, and unfortunately this argument does not help us argue against that.

Louis deRosset (2009) and Jessica Wilson (2010) both attack Hume's Principle while McPherson (2012) upholds it. All three authors agree that a strong version of Hume's Principle, defined in terms of identity, is not plausible. They agree that there are various kinds of counterexamples, including some of those I have mentioned. In different ways, they then aim to proceed with a weaker notion, one that concedes a certain range of counterexamples. But that range seems undefined, and it varies between the authors under discussion. DeRosset and Wilson think that some such weaker notion is still implausible while McPherson thinks that some such weaker principle is plausible.

(C) DeRosset argues against Hume's principle (deRosset 2009). He casts the principle in terms of things that are 'not wholly distinct', which includes parts and wholes, and he also includes sets and their members (deRosset 2009: 157–158). The last inclusion is surprising, since sets and members have incompatible properties in many cases, and hence must be wholly distinct if anything is. How can a non-spatio-temporal thing not be wholly distinct from a collection of spatio-temporal things? But (like Wilson) he wants to work with a principle that allows necessary relations between things related by the part-whole relation and the membership relation, and which, in that sense, are not 'wholly distinct'. Even so, he raises cases of essential origins as counterexamples to a Hume's Principle that is cast in terms of his notion of 'wholly distinct', since persons and their origins are necessarily connected things that are 'wholly distinct'. (That is, they are not related in the ways on his list.) Necessity of origin cases show that Hume's principle, thus formulated, is implausible. But 'wholly distinct' on this account seems to mean something like: not being identical, not sharing parts, not being constituttee and constituted, not being dependee and dependant, and not being determinates and determinables. But this is an unruly list. Strictly speaking, deRosset is right that the doctrine of the necessity of origin is not one of those relations. But the notions of 'wholly distinct', and the Hume's Principle that is constructed with that notion, has become a strange concoction.

(D) Wilson raises the cases of constitution, dependence, and determinables and determinates, as problems for Hume's Principle (Wilson 2010: 601–603). And she thinks that these relations do not hold between 'wholly distinct' facts. She then considers a weakened version of Hume's Principle that embraces these relations, which says that there are no necessary connections between things that are 'wholly distinct' in her constructed sense. She concludes that even this weaker principle has no intuitive support and the arguments for it are no good.

I suspect that I agree with both these authors, for the most part. But puzzles remain over Hume's Principle itself, which now seems to be a movable doctrine. Which relations constitute it and why?

(E) McPherson also wants to operate with a weak or ("modest") conception of Hume's Principle. Unlike Wilson and deRosset, he aims to retain such a principle, not reject it. He approaches this by deploying a notion of 'continuous' properties, which seems to include some of the cases I have adduced as counterexamples to Strong Hume's Principle. But McPherson's notion of 'continuousness' is somewhat opaque. It seems to include properties related by reduction as well as determinates and



determinables (McPherson 2012, 218, 227). He puts the latter case to one side as unproblematic. But it is a plain refutation of Hume's Principle if it is cast in terms of identity – the strong version. And if it is to be included in a wider notion, then, according to what principle? The resulting 'modest' Hume's principle, constructed with the notion of 'continuousness', is hard to apply. Even given what McPherson seems to want to include under it, there seem to be other cases of necessary connections between distinct facts or things beyond what he probably has in mind, such as cases of necessary origins. In one place he indicates what "continuous" might mean, when he says that a property is 'continuous' with natural properties when its "nature is to be understood in terms [...] that are themselves deeply naturalistic" (McPherson 2012, 207). This definition is epistemic, which is not what we want when sorting relations for some wider Hume's Principle. If we subtract the epistemic aspect, the relation just looks like reduction; if so, why did he not say that properties 'continuous' with natural properties are other natural properties? So 'continuous' properties are somehow similar? Anyway, this idea seems to be about the properties that are the relata, not the relation between them, and so it cannot help to formulate any version of Hume's Principle, where what we want is a non-arbitrary list of relations. McPherson did yeoman service by suggesting that Hume's Principle casts light on the supervenience objection(s) to non-naturalism. But the notion of 'continuousness', which is constitutive of McPherson's reconstructed Hume's Principle, does not help us. Certainly, it is not sharp enough to be used in an argument against moral nonnaturalism.

Philosophy in this area is rather confusing. There is a rag-bag of relations between two kinds of properties, which have no clear relation to each other, and with no obvious natural groupings. Furthermore, philosophers pick and choose among the relations – some they like and some they do not like – in a way that varies greatly and that seems to lack rationale. In the next and final sections, I attempt to impose order.

## **8. Diagnosis: Modal anti-realism and local-global essentialism**

Why is there anything to be said for Hume's Principle? To shake off any remnants of even *prima facie* attractiveness, and to see that it is accepting Hume's Principle that is *prima facie* implausible, consider that in its stronger identity form, it implies that there are only necessary relations between things, facts or properties that are identical. But that requirement, or consequence, is odd. Also odd is a weaker principle that implies that there are only necessary connections between things that are related either by identity or as parts and wholes.

What this, together with the tsunami of counterexamples, suggests is that Hume's Principle, in restricting necessities to identicals or mereologicals, is a dogmatic rejection of necessities altogether, or all except a trivial class of them. That rejection is much stranger than anything dreamed of in non-naturalistic moral realist metaphysics! The supervenience objection to moral realism depends on an implausible and revisionary doctrine about modality.

This diagnosis is the one given by McFetridge (1985). He claimed that there are connections between moral realism and modal realism. The supervenience objection to moral realism depends on Hume's Principle; but Hume's Principle is just a commitment to what we might call 'modal anti-realism'. Modal realism is best defined in terms of properties not in terms of objects, such as 'worlds' (see McGinn 1999 [1984] and McGinn 2000). We may take modal realism to be (roughly) the view that modal properties are genuine and judgement-independent properties of things. (If we take modal realism to be a doctrine about properties not objects, it turns out that David Lewis's objectual view is a modal anti-realist view; indeed, Lewis expresses reservations about the jargon of 'realism' in his Lewis 1986: viii, and he often describes himself as 'Humean'.) Now, if there are genuine mind-independent modal properties of things, then these properties bind one thing to another wholly distinct thing, or they bind one fact to another wholly distinct fact. They do not only bind one object to itself or one fact to itself.

Why believe modal realism? The usual argument is, first, that modal thinkers have a tacit commitment to mind-independent modal properties, plus, second, some kind of conservatism about our commitments given that the folk practice functions well and has benefits. But such conservatism about modal thought is flouted by the supervenience argument against moral non-naturalism.

Of course, someone sympathetic to Hume's general programme is likely to be against both non-natural moral properties and irreducible modal properties. But if one accepts non-natural moral properties, then one should

also accept that modal connections between them and their natural bases cannot be eliminated or reduced. One way to eliminate or reduce modal properties is to reduce them to identities. But the modal relations binding moral and natural properties go beyond what is allowed by Strong Hume's principle – that is, beyond necessities deriving from identities. Why on earth (or elsewhere) should identities be the only source of necessities?

That is a version of the diagnosis offered by McFetridge. But I want to pursue a more nuanced version of it, although I am sympathetic with McFetridge's more straightforward version. It might be complained, that given Hume's Principle, there remain some necessities, those underwritten by identity; so not all necessities are denied. However, it is certainly a major revision of our usual view of the extension of modal notions.

Is this a kind of stunted realism or anti-realism? I am not sure that there is a point coming down one side or the other. Suppose someone claims that there are only duties to ourselves and none towards other people: is that moral realism? Or anti-realism? Or what if it is claimed that moral facts are facts about what ideal observers would judge: realism or anti-realism? What about the view that scientific theoretical entities are only those that could be perceivable by equipment that enhances our seeing or hearing. Is that realism or anti-realism about theoretical entities? Or consider: social facts are facts about individual human beings. Realism or anti-realism? In each case, our folk theory about the extent of the phenomena is more or less respected.

Suppose that Hume's Principle is cast in the weak way – it only allows necessary connection between identicals and overlapping things. As we have seen, there are still many counterexamples. But the present question is whether saying that is a commitment to modal anti-realism. Yes, it is, if a modal realist is someone who is committed to retaining most of our pre-theoretic beliefs about the extension of mind-independent modal properties of objects or events. Many of these modal properties link things or events that are neither identical nor parts and wholes. Consider the necessity of origin. If we are to believe Saul Kripke (1980), there is a common-sense idea that in many cases, such as that of persons and their origins, one must come from the other. Yet they are wholly distinct. Accepting a Weak Hume's Principle therefore means rejecting the necessity of origin (as deRosset 2009 argued). Perhaps realism and anti-realism come in degrees, varying with the extent of the departure from folk theory or common sense. Matching pretheoretic extension would be a central aspect of this. (The case of zero extension, where no pre-theoretic doctrines are sustained, would be the extreme case of this continuum.) Or perhaps a theory could accept the pre-theoretic extension of a concept but reject some folk principles that govern thought of that kind. That would signal a retreat from full-blooded realism, just as much as a revision of the pre-theoretic extension.

We could define weak Hume's Principle to be an open family of views where modality is restricted, not to identity, but to some underspecified list of other relations: perhaps constitution, part-whole, determinate-determinable, as well as identity. But, I have protested, this is a chaos of relations! It is confusing to call Hume's Principle "Weak", when it is defined in terms of identity or mereology or some other relations. "Vague Hume's Principle" would be a better label.

Nevertheless, we need to understand why quite a few authors have had an urge to embrace some more expansive but vague version of Weak Hume's Principle. In moral philosophy, the question seems to be whether we can assimilate the necessity of M-N relations to some other cases, such as the determinate-determinable case, which for some reason, some philosophers (and this seems to vary quite a lot) find more comfortable.

But why go that route? What is the point of doing that? We start with some inventory of relations that some philosophers feel comfortable with. (These are phenomenological/sociological facts about philosophers.) Those philosophers, with their list of relations that they find cosy, look to see if M-N necessities fall under one of the cosy-feeling relations. But I ask: why does the M-N necessity relation somehow have to prove itself unlike, say, the determinable-determinate relation or the identity relation?

We can extract a diagnosis from this syndrome. Consider the necessity of personal origins. The property of personhood dictates modal connections, such as a connection to origins. Similarly, normative properties dictate a pattern of modal relations, those that moral supervenience principles enunciate. What emerges, I suggest, is this – and this is different from McFetridge's diagnosis: there are, broadly, two kinds of explanations of necessities. First, there are appeals to general metaphysical relations, which are 'topic neutral', such as

identity, the partwhole or determinate-determinable relations. Some necessities can be explained as instances of such general relations, which are often assumed to be unproblematic. Second, there are appeals to the essences of specific kinds of entities or properties, such as persons, normativity, scarlet, numbers or sets. These entities or properties dictate a range of necessary connections that just concern entities or properties of that type. Both such global and local explanations are essentialist: one appeals to the essences of identity, part-whole, and so on, while the other appeals to the essences of persons, colours, numbers, sets or whatever. Those who worry about moral-natural supervenience relations for non-naturalism are those who prefer global to local essentialist explanations of modal facts. But I cannot see what would justify that preference. We should expect there to be both global plus local essences.

In fact, the identity relation is not as innocent as we have been led to believe. (Hume says little about it.) That relation has an essence, which explains why Leibniz's Law holds as a modal principle governing identity:  $X=Y$  explains the fact that necessarily  $\forall x (Fx \leftrightarrow Fy)$ . (I do not agree with Colin McGinn's claim in his 2000 that Leibniz's Law assumes the identity of properties; instead, the necessary biconditionals are explained by the essence of identity. Perhaps, but why?)

The right view of the matter, I believe, is this: we should embrace both global and local essentialist commitments. There are general relations that have essences that dictate modal principles that apply to any kind of objects or properties that fall under them; and there are also specific kinds of objects or properties with essences that dictate necessities just for those kinds of objects and properties. Both are good, and neither is better than the other.

## 9. Coda

In either the Weak or Strong forms, Hume's Principles concerning modality are highly revisionary of ordinary modal thought. The supervenience argument against moral realism depends on a revisionary modal view. Insofar as the objection depends on an implausible modal view, the supervenience objection to nonnaturalist moral realism lacks force. Indeed, it looks as if the metaphysically bizarre view is not moral non-naturalism but the modal revisionary views assumed by the argument against it. If so, there is no reason to think that the supervenience commitments of non-naturalist moral realism are even slightly problematic. Hume's Principles, Strong and Weak, are dead. And with them is the parasitic metaphysical supervenience argument against non-naturalist moral realism. They can be buried in the same grave.\*

\*Many thanks to Tristram McPherson for comments on an earlier draft. Thanks also to questions from audiences at Reading University, Osaka University, and the Huazhong University of Science and Technology in Wuhan, and also to three referees for this journal. I note that Ian McFetridge was my doctoral supervisor

## References

- BLACKBURN, S. 1993 [1985], "Supervenience Revisited", in: *Essays on Quasi-realism*, Oxford: Oxford University Press.
- DEROSSET, L. 2009, "Production and Necessity", *Philosophical Review*, 118, pp. 153–181.
- ENOCH, D. 2011, *Taking Morality Seriously*, Oxford: Oxford University Press.
- FIELD, H. 2009, "Epistemology without Metaphysics", *Philosophical Studies*, 143, pp. 249–290.
- FODOR, J. 1981, *Representations*, Cambridge, MA: MIT Press.
- FODOR, J. 1987, *Psychosemantics*, Cambridge, MA: MIT Press.
- FODOR, J. and LEPORE, E. 1992, *Holism: A Shopper's Guide*, Oxford: Blackwell.
- HUME, D. 2000, *A Treatise on Human Nature*, Oxford: Oxford University Press. Q2
- HUME, D. 2002, *Enquiry Concerning Human Understanding*, Oxford: Oxford University Press.
- KIM, J. 1998, *Mind in a Physical World: An Essay on the Mind-Body Problem and Mental Causation*, Q3 Cambridge, MA: MIT Press.
- KRIPKE, S. 1980, *Naming and Necessity*, Oxford: Oxford University Press.
- LEWIS, D. 1986, *On the Plurality of Worlds*, Oxford: Blackwell.
- MACKIE, J. L. 1977, *Ethics*, Harmondsworth: Penguin.
- MCFETRIDGE, I. 1985, "Supervenience, Realism, Necessity", *The Philosophical Quarterly*, 35, pp. 245–258.
- MCGINN, C. 1999 [1984], "Modal Reality", in: *Knowledge and Reality: Selected Essays*, Oxford: Oxford University Press.
- MCGINN, C. 2000, *Logical Properties*, Oxford: Oxford University Press.

- MCPHERSON, T. 2012, "Ethical Non-Naturalism and the Metaphysics of Supervenience", in: R. Shafer-Landau, ed., *Oxford Studies in Metaethics Volume 7*, Oxford: Oxford University Press, pp. 205–234.
- SALMON, N. 1981, *Reference and Essence*, Princeton, NJ: Princeton University Press.
- WIELENBERG, E. 2014, *Robust Ethics: The Metaphysics and Epistemology of Godless Normative Realism*, Oxford: Oxford University Press.
- WILSON, J. 2010, "What is Hume's Dictum, and Why Believe It?" *Philosophy and Phenomenological Research*, 80, pp. 595–637.
- WITTGENSTEIN, L. 1922, *Tractatus Logico-Philosophicus*, London: Routledge.
- WITTGENSTEIN, L. 1979, *Wittgenstein and the Vienna Circle: Conversations Recorded by Friedrich Waismann*, ed. by Brian McGuinness, trans. by Joachim Schulte and Brian McGuinness, Oxford: Basil Blackwell.
- ZANGWILL, N. 1996, "Moral Supervenience", *Midwest Studies in Philosophy*, 20, pp. 240–262.
- ZANGWILL, N. 2005, "Moore, Morality, Supervenience, Essence, Epistemology", *American Philosophical Quarterly*, 42, pp. 125–130.
- ZANGWILL, N. 2006, "Moral Epistemology and the Because Constraint", in: J. Dreier, ed., *Contemporary Debates in Moral Theory*, Oxford: Blackwell, pp. 263–281.
- ZANGWILL, N. 2008, "Moral Dependence", in: R. Shafer-Landau, ed., *Oxford Studies in Metaethics Volume 3*, pp. 109–127.
- ZANGWILL, N. 2010, "Normativity and the Metaphysics of Mind", *Australasian Journal of Philosophy*, 88, pp. 21–39.
- ZANGWILL, N. 2011, "Negative Properties", *Noûs*, 43, pp. 528–556.
- ZANGWILL, N. 2013, "Knowledge and Truth", *Acta Analytica*, 28, pp. 139–144.
- ZANGWILL, N. 2017a, "Moral Dependence and Natural Properties", *Aristotelian Society Supplementary Volume*, 91, pp. 221–243.
- ZANGWILL, N. 2017b, "Epistemic/Non-epistemic Dependence", *Noûs*, <https://doi.org/10.1111/nous.12197>.