

The Supervenience Challenge to Non-Naturalism*

Pekka Väyrynen

p.vayrynen@leeds.ac.uk

Introduction

It is impossible that one action is morally impermissible and another permissible unless they differ also in some other respect; perhaps one involves stealing but the other doesn't. One person cannot be morally better or more virtuous than another without there being some other difference between them, such as that one is more reliably disposed to help others or to keep her promises. It cannot be that you have a reason to run for the bus on one occasion but no reason to do so on another if the situations differ in no other qualitative respect, such as how efficient each action will be for achieving your aims or for doing something good. If two persons are qualitatively exactly alike (or "indiscernible") in all other ways, they cannot but be morally alike as well.

The above claims look immensely plausible. Each is a way of saying that some normative feature (moral permissibility, moral goodness, reasons for action) is *supervenient*. Collectively they suggest that normative features, as a family comprising moral features and all other normative and evaluative features, such as aesthetic and prudential values and reasons – are supervenient. There are several different relations that go by the name of supervenience, but they all share this core idea: things cannot differ in one respect

*Final draft, 7 January 2016. To appear in *The Routledge Handbook of Metaethics*, ed. D. Plunkett and T. McPherson.

without differing in some other respect. (See McLaughlin and Bennett 2011 and McPherson 2015 for excellent overviews of supervenience in general and in ethics, respectively.) Because supervenience is a necessary connection, it requires explanation. Different metaethical theories may explain the connection in different ways. But if a theory cannot provide a good explanation, this makes it in one respect worse than its rivals. The supervenience challenge says this is the predicament of non-naturalist moral realism.

Non-naturalism is a form of *moral realism* – a family of views according to which there are moral truths that are objective and metaphysically robust. (If a proposition *p* objectively true, in the sense realists have in mind, *p* is true independently of anyone’s attitudes, stances, beliefs, and theories concerning whether *p*. Moreover, the realist notion of truth isn’t merely deflationary but carries ontological commitment.) In slightly different terminology, moral predicates, such as ‘good’, ‘wrong’, and ‘permissible’, refer to objective moral properties in a metaphysically significant sense, and at least some such properties have instances. If the claim ‘Stealing simply for one’s own gain is wrong’ is true, what makes it true is the fact that stealing simply for one’s own gain has a metaphysically significant attitude-independent property of being wrong.

The distinctive claim of *non-naturalist* moral realism is that these moral properties, and perhaps normative properties in general, are *sui generis* – significantly different in kind from any other properties. (Contemporary non-naturalists include Hampton 1998, Shafer-Landau 2003, Cuneo 2007, FitzPatrick 2008, Enoch 2011, Wielenberg 2014, Cuneo and Shafer-Landau 2014, and Skarsaune 2015. Some of these authors characterize non-naturalism as “robust” moral realism.) To spell this out a bit, the non-naturalist thinks that at least some normative properties aren’t identical with any natural or supernatural properties, nor do they have a real definition, metaphysical reduction, or any other such tight metaphysical explanation wholly in terms of natural or supernatural properties. Normative properties are, in short, *discontinuous* with natural and supernatural properties. Supernatural properties are tricky to distinguish from the non-natural, but are meant to include non-normative properties such as being willed by God. Natural properties are

(i) such that any synthetic propositions about their instantiations (including propositions stating any empirical regularities in which they could figure, or are “such as to” figure) are empirically defeasible or (ii) reducible to properties that satisfy (i) (Smith 2000: 211-12; Copp 2003: 185). The properties, kinds, and causal systems studied by the natural sciences, psychology, and at least some social sciences come out as natural on this criterion, but so can properties to which we have reliable non-scientific empirical access. Non-naturalists deny that normative properties are anything like that. What makes them so? Perhaps, as many non-naturalists suggest, that they are *irreducibly* normative. (That would distinguish normative properties also from mathematical properties, which might be neither natural nor supernatural.) Be that as it may, the core idea of non-naturalist moral realism is that moral (and other normative) properties are metaphysically *sui generis* and at least some of these properties have instances.

So understood, non-naturalism and the supervenience of the moral jointly entail that moral properties supervene on some properties with which they are discontinuous. The supervenience challenge says that non-naturalists cannot explain this connection without making commitments that count significantly against their view. So, unless non-naturalism has other merits worth the cost, we should reject it. In what follows I formulate a version of this challenge more carefully, consider the most promising non-naturalist replies to it, and suggest that none of the replies is as yet fully effective.

The Supervenience Challenge

Virtually all metaethical theories seek to accommodate in some way the idea that there can be no moral difference without some other difference. It is less clear whether any particular way of fleshing out this core idea of supervenience is similarly close to common ground in metaethics. Perhaps there is no supervenience claim that does serious argumentative work without begging any important metaethical questions (Sturgeon 2009). But some claims might be suitably neutral for a particular philosophical purpose. I'll first iden-

tify a supervenience claim that isn't question-begging in the context of the supervenience challenge to non-naturalism, and then explain the challenge.

If things cannot differ morally without differing in some other qualitative respect, this doesn't seem to be merely a claim about how things must be in a given world. To use a classic example from R. M. Hare, if St. Francis was a good person, then *anyone* exactly like him in all other respects *couldn't but have been* good as well (Hare 1952: 145). What difference would it make whether a duplicate of St. Francis were actual or merely possible? None, it seems. Accordingly, the supervenience challenge to non-naturalism is best understood in terms of *strong* supervenience, which constrains variations among any possible entities. (The contrast is with *weak* supervenience which only constrains entities in the same possible world; see Kim 1984.) Strong supervenience, like supervenience in general, is a *purely modal* relation. It only entails that certain patterns of variation hold between how things are in one respect and some other respect. It doesn't follow that the supervenient properties are metaphysically dependent on or explained by properties in the supervenience base. (Everything supervenes on itself, and the supervenience of one family of properties on another is compatible with the supervenience of the latter on the former.)

In what follows, it'll help to be clear about the formal structure of strong supervenience claims, in particular that they invoke necessity twice:

$$(SS) \quad \Box(\forall F \text{ in } \alpha)(\forall x)[Fx \rightarrow (\exists G \text{ in } \beta)(Gx \ \& \ \Box(\forall y)(Gy \rightarrow Fy))]$$

This formula requires both explanation in ordinary language and interpretations of the schematic variables and the necessity operators (marked by '□'). Variables 'x' and 'y' pick out individuals like persons or actions. Thus (SS) isn't a *global supervenience* claim to the effect that the *world* couldn't have been different in one type of respect without some other type of difference, but an *individual supervenience* claim. (For complications regarding strong global supervenience, see McLaughlin and Bennett 2011: §4.4.) I'll take 'α' and 'β' to pick out families of properties. We'll be interested in the case where α is the family of moral properties; so 'F' stands for some specific moral property in that family. (Of β more below.) So interpreted, (SS) states an *ontological*

connection between families of *properties*, not an *ascriptive* connection between types of *judgments* (Klagge 1988). This interpretation isn't available to metaethical expressivists, who eschew ontological commitment to moral properties and treat supervenience as a constraint on normative judgments (Hare 1952: 80-81; Hare 1981; Blackburn 1985). (Complications that arise in the context of "quasi-realist" expressivism are usefully explored in Dreier 2015.)

The interpretation of the necessity operators in (SS) taken as a moral supervenience claim is controversial. The innermost necessity is typically taken as metaphysical: whenever something has a moral property M, it has some (possibly very complex) property that metaphysically necessitates M. (We'll revisit this assumption in the last section.) This interpretation reflects the widely held view that the basic principles of morality are metaphysically necessary. The outermost necessity is most often taken as conceptual, but sometimes as metaphysical. (For our purposes, conceptual necessity may be understood as metaphysical necessity knowable by conceptual reflection. In other contexts this might require finessing. Taking the outermost necessity as metaphysical is compatible with but not entailed by the view that moral supervenience is a substantive *moral* truth; see Kramer 2009: ch. 10.)

Given this interpretation, (SS) says the following: as a matter of conceptual/metaphysical necessity (the outermost '□'), when something has a moral property, it has some (possibly complex) property (from class β) such that, as a matter of metaphysical necessity (the innermost '□'), anything that has the latter property has the moral property. Closer to ordinary English, we might say that absolutely any morally wrong action has some features such that anything else with those features cannot but be wrong as well, and likewise for any other moral property.

In what sort of respect must things differ in order for moral differences to be possible? Many interpretations of ' β ' for moral supervenience have been proposed: the factual, the natural, the descriptive, the non-moral. These may all be different, and each interpretation comes with certain costs to metaethical neutrality (Sturgeon 2009). Fortunately there is a way forward. Call a property *morally involving* if either it is a *sui generis* moral property or its

correct analysis ineliminably mentions such properties, and say that a *base property* is any property that isn't morally involving (McPherson 2012: 213-4). This specification of the "supervenience base" generates the following strong supervenience claim:

SUPERVENIENCE Necessarily, when something has any moral property, it has some base property such that, as a matter of metaphysical necessity, anything that is exactly alike it with respect to the base property also has the moral property.

SUPERVENIENCE is a *general* rather than *specific* supervenience claim: it only requires *some or other* difference in base respects for a moral difference, and says nothing about *which* specific base properties moral properties supervene on. Thus it is neutral between various first-order theories in normative ethics.

Nor does SUPERVENIENCE beg any important metaethical questions at stake in the supervenience challenge to non-naturalism. If moral properties are both supervenient and *sui generis*, as non-naturalists typically claim, they must supervene on a set of base properties with which they are discontinuous. So SUPERVENIENCE follows from non-naturalism. (When talking about non-naturalism, it is usually safe to assume that all base properties are non-moral properties, and I'll sometimes talk this way.) Moral realists in general can accept SUPERVENIENCE. Naturalist moral realism, for example, says that moral properties belong to the category of natural properties. This entails that moral properties are base properties, not morally involving properties. Thus, since everything supervenes on itself, moral properties cannot change without some base properties also changing, namely themselves. So SUPERVENIENCE follows trivially from moral naturalism. Nihilist metaethical views can also accept SUPERVENIENCE. For if nothing instantiates moral properties, it follows trivially that any items that are alike with respect to all base properties are exactly alike morally. They'll be morally void.

The impossibility of moral differences without base differences is plausibly not brute, so it requires explanation. The version of the supervenience challenge to non-naturalism which I'll discuss is generated by conjoining SUPERVENIENCE with a claim concerning what non-naturalism says about SU-

PERVENIENCE and a plausible methodological assumption. (My formulation of the challenge will largely follow McPherson 2012: 217-9.)

The methodological assumption in question concerns necessary connections between discontinuous properties. Such connections often have explanations. Necessarily, if something is a brick, it is identical to itself. But this necessary connection between the seemingly very different kinds of properties of being a brick and being self-identical has an explanation: everything is necessarily identical to itself and any conditional with a metaphysically necessary consequent is itself metaphysically necessary (Leary forthcoming). Since many necessary connections between discontinuous properties have explanations, leaving such a connection brute and unexplained looks like a cost. One methodological principle that one might take this point to suggest is the following:

MODEST HUMEAN Commitment to brute necessary connections
between discontinuous properties counts significantly against
a view. (McPherson 2012: 217.)

MODEST HUMEAN allows that positing brute necessary connections might not rule out a view if it has merits that overall outweigh the cost (McPherson 2012: 218.) Thus MODEST HUMEAN is weaker than the more familiar (and more controversial) “Hume’s Dictum” which rules out any metaphysically necessary connections (brute or otherwise) between distinct entities (Lewis 1983: 366; for critical discussion, see Wilson 2010). The supervenience challenge might also get by with something weaker still than MODEST HUMEAN (Leary forthcoming).

Since the non-naturalist claims that moral properties are *sui generis*, she cannot explain the necessary connection in SUPERVENIENCE by the usual expedients of analysis, reduction, or identity. Such explanations would make moral properties continuous with base properties. But if moral and base properties are discontinuous, why should it be impossible for things to differ morally without differing in base respects? Whence a bar on such variations? The following claim about moral non-naturalism has at least *prima facie* plausibility:

BRUTE CONNECTION The non-naturalist must take the supervenience of moral properties on base properties to involve a brute necessary connection between discontinuous properties. (McPherson 2012: 217.)

SUPERVENIENCE, BRUTE CONNECTION, and MODEST HUMEAN jointly entail that moral non-naturalism is committed to brute necessary connections which count significantly against its plausibility.

This formulation of the supervenience challenge to moral non-naturalism differs from others that wear the same label. Perhaps the most discussed version of the challenge is due to Simon Blackburn, who takes a claim much like SUPERVENIENCE to be a *conceptual* truth and uses its alleged conceptual status to argue against moral realism in general (Blackburn 1984: 182-90; 1985). This is an overreach: supervenience doesn't raise problems for naturalist moral realism (Dreier 1992; Sturgeon 2009). The supervenience challenge above targets only moral non-naturalism. Moreover, it can be stated by treating the outermost necessity in SUPERVENIENCE as metaphysical, bracketing the question whether that necessity might also be conceptual.

I'll now turn to the most promising replies to this version of the supervenience challenge. I'll first consider the prospects of rejecting BRUTE CONNECTION and then discuss whether non-naturalists might avoid the challenge by rejecting SUPERVENIENCE. I'll largely bracket MODEST HUMEAN since it isn't special to ethics but a general methodological principle that requires broader assessment. (For a *prima facie* case that MODEST HUMEAN is self-undermining, see Wielenberg 2014: 33-34.)

Reject BRUTE CONNECTION?

Among the three premises that generate the supervenience challenge to non-naturalism, BRUTE CONNECTION seems the most vulnerable. Even if SUPERVENIENCE isn't explicable by some continuity between moral and base properties, this doesn't entail that it has no other explanation. Non-naturalists may thus seek to reject BRUTE CONNECTION by offering a positive explana-

tion of the kind of necessary connection between discontinuous properties which they take SUPERVENIENCE to assert.

The first response to BRUTE CONNECTION I'll discuss is the Conceptual Strategy. Many philosophers regard SUPERVENIENCE as a conceptual truth: flouting it manifests a conceptual deficiency. The Conceptual Strategy is thus to explain SUPERVENIENCE by the nature of moral concepts in general. We would be deeply puzzled by people who, for instance, regard St. Francis as a good person but at the same time think that there might have been another person with exactly the same character and behavior and placed in exactly the same circumstances, but differed from St. Francis only in not being a good person. Our puzzlement is evidence that such speakers either manifest a deficient grasp of the concept of moral goodness or are talking about something else, and thus evidence that SUPERVENIENCE is a conceptual truth. No further explanation is needed. Explaining general moral supervenience saddles non-naturalism with no additional ontological cost beyond its distinctive commitment to *sui generis* moral properties. (Stratton-Lake and Hooker 2006: 164; Enoch 2011: 149; Olson 2014: 96-99; Cuneo and Shafer-Landau 2014: 429-30.)

The Conceptual Strategy fails, for two reasons. The first is that it fails to provide a right kind of metaphysical basis for the connection which it represents as conceptually necessary. Not all analytic truths have a metaphysical basis; for example, 'I am here now' is analytic but not metaphysically necessary. (I owe this point to Jamie Dreier.) But many do. Even if (as some philosophers think) it is a conceptual necessity that magnets attract iron, it still remains to be explained why or how magnets do this. For realists about magnets and iron, it is fundamentally the world that guarantees attraction between magnets and iron, not our representations of them. The missing explanation will presumably be the same physical explanation that is accepted by people who don't assume that the necessity by which magnets attract iron is conceptual (Sturgeon 1999: 95). SUPERVENIENCE looks more like 'Magnets attract iron' than 'I am here now' in this respect, if only because it is supposed to concern a necessity irrespective of whether the necessity is conceptual. And since SUPERVENIENCE is a connection between families of prop-

erties, presumably it holds (if it does) irrespective of whether it is reflected in our concepts. (It should hold irrespective of whether a person who thinks that someone just like you in every non-moral respect could have been a much worse person is conceptually deficient or just bad at moralizing.) Thus it is hard to see how, at least given the terms and conditions of moral realism, the conceptual status of SUPERVENIENCE is supposed to explain the necessary connection it states between moral and base properties. The explanation will point to a metaphysical connection which according to this reply is *also* reflected in moral concepts. The question then arises whether that necessary metaphysical connection is brute.

The second problem with the Conceptual Strategy is that it explains the wrong necessity. (This objection is due to Dreier MS.) Recall that strong supervenience claims like SUPERVENIENCE contain two necessity operators: *necessarily*, if any two possible individuals are alike in every base respect, then they *must* be alike in every moral respect. Now, when proponents of the Conceptual Strategy say that SUPERVENIENCE is conceptually necessary, they mean that the outermost necessity is conceptual. But what about the innermost one? Reading it as conceptual would entail that the base properties (whatever they are) attach to the supervening properties with conceptual necessity. But that isn't true in the moral case. Propositions that state connections between particular non-moral and moral properties are synthetic, not analytic. That is why the innermost necessity operator in SUPERVENIENCE is typically interpreted as metaphysical. So the claim which the Conceptual Strategy represents as a conceptual truth is a truth about a metaphysical necessity. But what the supervenience challenge to non-naturalism represents as needing explanation is the metaphysical necessity by which base properties necessitate moral properties. The Conceptual Strategy thus explains the wrong necessity. The puzzle for non-naturalism remains even if SUPERVENIENCE is a conceptual truth: if moral properties are *sui generis*, how can the necessary metaphysical connection between non-moral base properties and moral properties (marked by the innermost necessity operator) be anything but brute?

A natural alternative to the Conceptual Strategy is to seek a metaphysical explanation of SUPERVENIENCE – a Metaphysical Strategy. Think of the challenge this way: how could there be properties which are *sui generis* and yet supervene in the way SUPERVENIENCE says? Any necessary connection to which non-naturalists may appeal in giving an explanation must hold between discontinuous properties, because otherwise the explanation won't be non-naturalist. And that connection must have an explanation, because explaining one necessary connection by relying on another brute necessary connection between discontinuous properties would merely relocate the problem. The Metaphysical Strategy seeks to identify a necessary metaphysical connection that avoids these problems.

One proposal here is to take moral facts to be *exhaustively constituted* by non-moral facts (Shafer-Landau 2003: 87). Suppose that the fact that Jane is generous is exhaustively constituted by non-moral facts concerning her disposition to assist those in need without a motive of self-interest, and the like (Shafer-Landau 2003: 75). If other instances of generosity can be exhaustively constituted by different non-moral facts, then generosity isn't reducible or definable by any particular non-moral constitution base. And yet if facts to the effect that someone is generous are always exhaustively constituted by non-moral facts, then things cannot differ with respect to their generosity without differing non-morally. Thus is SUPERVENIENCE explained.

The constitution view fails to help non-naturalists, for two reasons. First, if each instance of a moral property is wholly constituted by some concatenation of non-moral properties, this threatens to make moral properties continuous with non-moral properties. The view that instances of mental properties are constituted by physical properties tends to count as a form of physicalism. And consider material constitution: a statue might have modal properties that the clay constituting it lacks without thereby being a *sui generis* type of entity. (Clay-arranged-statue-wise is still clay.) Second, explanations that are murkier than what they aim to explain are (all else equal) no good. Suppose a fact is an instantiation of a property by an object (at a time). It seems that only the property component of facts can carry the structure that fact constitution would require. But the claim that one and the same *prop-*

erty, F, is constituted by G, H, and I on one instantiation but by J, K, and M on another looks murky – not because it is difficult to grasp (as explanations in quantum physics might be to a lay person), but because its coherence is dubious. How can the constitution of a type entity vary across its tokens in this way? (This problem doesn't arise in the case of the material constitution of physical objects.) No murky claims about fact constitution are required if we say instead that one object can have F in virtue of having G, H, and I and another object can have F in virtue of having J, K, and M. For instance, saying that being painful and being autonomy-undermining can each make an experience bad commits us to no particular metaphysics of badness. (Ridge 2007 discusses a possible reply that goes beyond our scope: if moral properties are understood as *tropes*, or abstract particulars, then what gets constituted on different occasions is a qualitatively similar but numerically distinct entity.)

Nearby variants of the Metaphysical Strategy look to fare no better. Suppose we say instead that moral properties are *realized* by non-moral properties, on analogy with how mental states such as beliefs or pains can be realized by different physical substrata (Shafer-Landau 2003: 77). This analogy is dubious even if we can make sense of moral properties as a kind of functional properties. If a subject's beliefs are realized (at a given time) by brain states, this excludes the realization of her beliefs (at that time) also by silicon-based states or computer hardware. But one can be virtuous by manifesting various different combinations of wisdom, generosity, courage, and the like. This cannot be explained by saying that the latter are realizers of virtue. Analogies between moral non-naturalism and non-reductive physicalism in the philosophy of mind seem generally suspect (Ridge 2007; McPherson 2012: 224-27).

The Metaphysical Strategy isn't fully exhausted by these options. One might, for instance, think that it is in the nature or essence of moral properties to satisfy SUPERVENIENCE in some specific way or other (Wedgwood 2007: 151, 207). Is this necessary connection between (some or other) base properties and the nature of moral properties itself a brute connection between discontinuous properties, and would it be objectionably brute? This is difficult to assess in abstraction from concrete proposals about the nature

of moral properties. Thus I'll merely note the possibility of this kind of explanation of SUPERVENIENCE. (Not all of the necessities involved with SUPERVENIENCE might be explicable this way; see Wedgwood 1999 and 2007: 207-20. Wedgwood argues that the remaining necessities involving specific supervenience facts are explicable by appeal to contingent facts. Assessing this account would require a long foray into modal logic; but see Schmitt and Schroeder 2011.)

The last response to BRUTE CONNECTION which I'll consider is the Substitute Strategy. The idea is that the non-naturalist isn't committed to necessary connections between discontinuous properties, but only to certain tolerably brute necessary normative truths. The most developed instance of the Substitute Strategy is T. M. Scanlon's explanation of normative supervenience. (Scanlon thinks that normative facts exist in an ontologically lightweight sense, so he isn't a "robust" non-naturalist realist. But the view below is compatible with non-naturalism.) Consider ordinary statements of reasons for action, such as that "the fact that the edge of a piece of metal is sharp is a reason for me, now, not to press my hand against it" (Scanlon 2014: 30). This ordinary reason relation only obtains when the piece of metal is in fact sharp, and it is normatively "mixed" thanks to its non-normative element. But the mixture has an essentially normative component, of the form $R(p, x, c, a)$, where R relates a proposition p , an agent x , a circumstance c , and an action a . To get a rough intuitive grip, think of statements of the form $R(p, x, c, a)$ as saying that "if p were true, and x were in circumstance c , then one reason for x to do a would be that p " (Schroeder 2015: 196).

These R truths are normatively "pure" in the sense that they hold irrespective of how the non-normative facts are (Scanlon 2014: 37-38). By contrast, ordinary reason claims like the one concerning my reason not to press the metal are contingent, since it is only contingent that the metal is sharp. What allows Scanlon to use the special normative relation R to explain normative supervenience is his view that $R(p, x, c, a)$ is *always* necessary, if true in the first place (Scanlon 2014: 40-41). Provided that the agent or the circumstances are made sufficiently specific, changing the agent or circumstances in any way delivers a different tuple from $\langle p, x, c, a \rangle$. Tuples that stand in R

thus have their non-normative features necessarily. So when R holds of $\langle p, x, c, a \rangle$, it does so necessarily, no matter how the non-normative facts had been. (We'll return to similar ideas in the last section.) But although pure normative truths of the form $R(p, x, c, a)$ are necessary, they bear no (other) necessary connection, brute or otherwise, to non-normative facts. So they generate no commitment to BRUTE CONNECTION. It is only the mixed normative truths that cannot vary without a non-normative difference. But the supervenience of mixed normative facts on the non-normative is explained by the way they are partly constituted by a relationship to their constituent non-normative facts. So mixed normative facts generate no commitment to BRUTE CONNECTION either. In sum, SUPERVENIENCE is explained on the basis of a conception of "pure" normative facts as facts about a special relation $R(p, x, c, a)$ such that, if they are necessary, then ordinary "mixed" normative facts supervene on the non-normative facts. (Schroeder 2014 and Skarsaune 2015 also offer to non-naturalists an explanation of supervenience based on a conception of basic normative truths on which such truths are always necessary.)

Unfortunately the Substitute Strategy avoids brute necessary relationships between discontinuous properties only by replacing them with other sort of brute necessary relationships in the vicinity. As Mark Schroeder notes, "to say that R ever holds of any tuple $\langle p, x, c, a \rangle$ is just to say that there are some necessary relationships that hold among wholly distinct entities" (Schroeder 2015: 197). Is the distinction between brute necessary truths and brute necessary connections really robust enough to help the Substitute Strategy avoid the kinds of theoretical costs that motivate MODEST HUMEAN?

In this vicinity lurks also a more general worry about treating normative truths as brute, which needn't appeal to MODEST HUMEAN. The Substitute Strategy aims to make it less puzzling how two discontinuous types of fact could be linked as tightly as SUPERVENIENCE says by representing the linkage between normative and non-normative facts as itself a set of normative truths, of the form $R(p, x, c, a)$. This explains the general supervenience of normative properties on non-normative properties, whatever they are, in terms of a set of particular pure normative truths which delivers a corre-

sponding set of truths about which specific non-normative facts are linked with which specific mixed normative facts. (The issue here isn't whether the pure normative content of the R truths can be understood in other terms, by giving a reductive definition of R. The issue is what explains why R holds of some tuple when it does.) But for each pure normative truth we can ask: why does R hold of $\langle p, x, c, a \rangle$, not of some other tuple $\langle q, y, d, b \rangle$? The distribution of the R relationship over facts, agents, circumstances, and actions shouldn't be arbitrary. The same question extends to the necessity of these relationships: when $\langle p, x, c, a \rangle$ stands in R, why couldn't that tuple have failed to do so? When things are claimed to have some property necessarily, a demand for explanation is usually legitimate. (As the logically weaker demand, contingency is the default status for truths, at least when their modal status is open to dispute. And the view that there are no unexplained necessities is a serious contender in metaphysics.) Why should the R facts be an exception? Yet it seems to follow from Scanlon's view that the particular pure normative truths have no explanation (Scanlon 2014: 44). A theory that provides no explanation is in that respect worse than one that does, and in any case worse off insofar as the demand for explanation is legitimate. Those who find it legitimate are bound to regard Scanlon's explanation of SUPERVENIENCE as leaving too many significant normative truths unexplained. (Largely the same worries as above can be raised against explanations of specific supervenience relations according to which there are metaphysically necessary "normative laws" which specify that if something has certain base properties, then it has certain normative properties, but no explanation for why the normative laws are what they are; see e.g. Enoch 2011: 142-8.)

Reject SUPERVENIENCE?

A different way to defend moral non-naturalism against the supervenience challenge is to reject SUPERVENIENCE. This strategy might seem quixotic since SUPERVENIENCE seems to be supported by highly compelling intuitions. But if these intuitions could be captured otherwise, there would be room to reject

SUPERVENIENCE. Call this the Contingency Strategy. (The label is due to Dreier MS. If SUPERVENIENCE is false when the innermost necessity is read as metaphysical, then even the most basic moral principles are metaphysically contingent.) To warrant rejecting SUPERVENIENCE, the Contingency Strategy should give non-naturalism some distinctive explanatory advantage. But that is by no means clear.

The most developed form of the Contingency Strategy is due to Gideon Rosen (MS). Rosen argues that SUPERVENIENCE and non-naturalism form an inconsistent triad with “essentialism” about metaphysical modality, and that out of these three claims we should reject SUPERVENIENCE. Strong supervenience entails that for each moral property M there is a condition consisting of a set of base properties, Φ , which is equivalent to M as a matter of (metaphysical) necessity. At minimum Φ is a vast disjunction of the complete specifications of all metaphysically possible bearers of M in all of their base respects (Jackson 1998: 122-23). Rosen further supposes that Φ is a naturalistic condition specifiable in wholly non-normative terms. (Many non-naturalists agree, such as Shafer-Landau 2003.) But now assume the essentialist theory of metaphysical modality: for a proposition p to be metaphysically possible just is for p to be logically consistent with all of the essential truths, where an essential truth about a given item x (an object, property, relation, etc.) is a truth that obtains in virtue of x 's nature or identity (Fine 1994). (For instance, being human is one of the things that lie in the essence of Socrates, but being a member of the singleton set {Socrates} isn't.) It follows that for each moral property M , there is a non-normatively specified naturalistic condition Φ such that for some item x , the equivalence between M and Φ is necessary in virtue of x 's nature. Isn't this a form of moral naturalism? The necessary equivalence would be a synthetic truth, but Rosen takes a distinctive commitment of moral non-naturalism to be that “someone who knew the natural facts and the essences might still be in the dark about the synthetic principles that connect the normative facts to their non-normative grounds” (Rosen MS: 12). In short: SUPERVENIENCE and essentialism jointly rule out non-naturalism.

Why not put essentialism in the reject pile instead of SUPERVENIENCE? Essentialism requires very substantial metaphysical commitments regarding the essences of things and casts metaphysical necessity more narrowly than many philosophers think. If, however, you are happy to work with essentialism, now ask yourself: in virtue of what item's nature might the necessary equivalence between (say) moral rightness and Φ hold? It won't be anything about Φ . For if the properties in Φ are wholly non-normative, then their natures are normatively silent. Nor will it be anything about moral rightness. For rightness won't have heard of many of the non-normative properties and relations in Φ , whatever they may be. It is hard to see what other item's nature could rule out the possibility that something is wrong but satisfies Φ . (For discussion, see Leary forthcoming.) Given essentialism, the metaphysically necessary equivalence between Φ and moral rightness fails if it is logically consistent with all the essential truths that something is wrong but satisfies Φ . And if the necessary equivalence fails, SUPERVENIENCE is false. The point can be brought out with putative counterexamples:

Consider a world w that is just like the actual world in non-moral respects, but in which act utilitarianism is true. Your act of reading this paper, A , would have been wrong if w had been actual. No matter how much benefit the world derives from your reading this paper, you would have done more good licking stamps for Oxfam instead. So we have a world w in which $D(A)$, a description that gives a complete specification of the wholly non-moral features of A , is true and A is wrong. Together with the actual world – where $D(A)$ is true and A is not wrong – this yields a counterexample to supervenience. (Adapted from Rosen MS: 3.)

The point may also be put using an epistemological heuristic. You might know all there is to know about the properties and relations in $D(A)$ without knowing whether A is right or wrong. And you might know all there is to know about the nature of wrongness without knowing whether $D(A)$ specifies one of the wrong acts. So it seems that nothing in the essences of things makes w impossible.

The metaphysical contingency of even the most fundamental explanatory principles of ethics – candidates for which include act utilitarianism, Kant’s categorical imperative, Ross’s plurality of principles of *prima facie* duty, and so on – might seem a suspect result. If St. Francis was a good person, wouldn’t absolutely anyone exactly alike him in their non-moral properties have to have been good as well? Rosen offers an innovative explanation of these compelling intuitions. Although no moral truths are absolutely necessary, some are “fact-independent,” where p is fact-independent if p is the case and would have been the case no matter how things had been in wholly non-normative respects (Rosen MS: 16). (Cohen 2003 introduces a similar notion of the fact-independence of moral principles, but in a different context.) We might alternatively express this by saying that moral principles hold as a matter of a *sui generis* type of *normative necessity* which isn’t reducible to metaphysical or natural necessity (Fine 2002). (Rosen offers fact-independence as an explication of Fine’s notion of normative necessity. Scanlon suggests that the necessity of his pure normative facts is an instance of normative necessity; Scanlon 2014: 41 n. 40. See also Cuneo 2007b: 863-71.) On this view, act utilitarianism might still be true in some worlds even if it is false in the actual world and would have been false no matter how the non-normative facts had been. Thus it would be no objection to the moral principles governing your (hopefully permissible!) reading of this paper to say that those principles would have been false if act utilitarianism had been true. If a moral principle is fact-independent, it would have been true no matter what we had thought or done, no matter how hard we tried to falsify it, no matter what the laws of nature had been, and so on. These claims are analogous to the view that laws of nature aren’t metaphysically necessary. Those taking this view usually think that laws of nature are explained by some metaphysically contingent facts. Is anything similar true of fact-independent moral principles? What metaphysically contingent facts would explain the important modal claims which morality supports, such as that if torture is wrong, then torture would have been wrong no matter how the non-moral facts had been?

Be that as it may, the argument offers powerful reasons to regard fact-independence as an important feature of moral principles and to think that SUPERVENIENCE, understood as constraining all metaphysically possible individuals, is controversial in its own right and not required to make sense of the practice of moral theory (Rosen MS: 17-19). Perhaps nothing important is lost if we instead adopt the weaker supervenience thesis that moral properties supervene on base properties as a matter of normative necessity. (Non-naturalism will still be compatible with a restricted strong supervenience claim: it is metaphysically/conceptually necessary that if any items x and y in worlds governed by the same set of basic moral principles M^* are alike in all base respects, then x and y must be alike in all moral respects. Note that this would make supervenience a substantive moral truth; cf. Kramer 2009: ch. 10. It is less clear that nothing important is lost if non-naturalists instead follow Hills 2009 and give up SUPERVENIENCE for the view that moral differences are merely constantly conjoined with base differences.) I'll close with three reservations about the Contingency Strategy.

First, the Contingency Strategy won't help non-naturalism in particular because its characterization of moral naturalism is too narrow. The distinctive commitment of moral naturalism is that moral facts and properties belong to the category of natural facts and properties. It doesn't follow that a property counts as natural only if it has a necessary equivalent specifiable in wholly non-normative terms, since moral properties might meet the criterion of naturalness directly in their own right (Sturgeon 2003: 536-40). Naturalism secures SUPERVENIENCE trivially: moral properties cannot change without a change in some natural properties, namely themselves.

Second, the Contingency Strategy might leave something morally important unexplained. If act utilitarianism is the basic principle of morality in some possible worlds but not ours, why is that? If sexism, speciesism, or violations of other people's bodily integrity are moral obligations in some metaphysically possible worlds, how come isn't ours one of them? (Any particular examples will be controversial: the metaphysical contingency of moral principles doesn't imply that anything could have been morally permissible or required. But constraints on what moral principles can be like, such as

consistency and universalizability, may not suffice to rule out morally problematic variation. For instance, why cannot the moral judgments of a fanatic Nazi who thinks he himself ought to have been exterminated if he had been Jewish be consistent and universalizable? For discussion, see Hare 1981.) The Contingency Strategy seems bound to say that these questions have no answer. So if it is legitimate to demand answers to them, failure to explain normative necessities would still be a strike against a theory that posits them. Should no such explanation be forthcoming, the supervenience challenge to non-naturalism might retain some of its bite even if we rejected strong supervenience claims that constrain moral differences across all metaphysically possible worlds. Moreover, even if the brute metaphysical contingency of the basic moral principles weren't problematic in itself, their seemingly consequent brute *authority* might still be. Act utilitarianism might (as a matter of metaphysical possibility) just as well have been the basic principle of morality. So why *shouldn't* we be guided by it, in lieu of whatever moral principles are basic for our world? Those who want the basic principles of morality to close normative questions concerning their authority may therefore find the Contingency Strategy argument unsatisfactory. (One explication of such normative questions can be found in Korsgaard 1996.)

Third, the Contingency Strategy seems to entail too much moral luck. Given moral contingency, it is metaphysically possible that the world could have been just as it is in wholly non-moral respects but our actions had been profoundly morally wrong. For if the basic moral principles had been different, then what we regard as morally creditable might have been morally monstrous. Nor would we have known this: in any metaphysically possible world which is non-morally just like the actual world, we would have the same moral beliefs we actually have; in counter-moral worlds these beliefs would be badly mistaken. If the basic moral principles are fact-independent but metaphysically contingent, we are extremely fortunate that the things we regard as morally innocuous aren't systematically morally monstrous. But it is incredibly hard to believe that we are *merely lucky* that the concern and respect with which we try to treat others and the care and love with which we try to raise our children are morally commendable rather than unspeakably

and grotesquely evil. Yet that is our situation, according to the Contingency Strategy. (This argument is due to Dreier MS.)

In conclusion, the Contingency Strategy is intriguing and deserves further discussion both in its own right and as a response to the supervenience challenge to moral non-naturalism. As it stands, however, it seems to provide no distinctive explanatory advantage to non-naturalism or, therefore, to warrant the rejection of SUPERVENIENCE.

Conclusion

The aim of this paper was to formulate a compelling version of the supervenience challenge to non-naturalist moral realism and assess some of the most promising responses to it. I have focused on two main strategies: rejecting the relevant supervenience thesis (SUPERVENIENCE) and rejecting the claim that non-naturalists must take moral supervenience to involve brute necessary connections between discontinuous properties (BRUTE CONNECTION). The responses I have considered raise fascinating issues. Some of these responses are nonetheless ineffective and even the most promising attempts suffer from some explanatory shortcomings. The demands for explanation on which these shortcomings are premised aren't themselves uncontroversial, however. Nor did I discuss how well some of the responses might work together if combined. The supervenience challenge to non-naturalism is thus not a closed chapter in metaethics.

Acknowledgments

Thanks to John G. Bennett, Tristram McPherson, David Plunkett, Lea Schroeder, Knut Olav Skarsaune, and Mark van Roojen for comments on earlier drafts.

References

- Blackburn, S. (1984) *Spreading the Word*, Oxford: Oxford University Press.
- (1985) “Supervenience Revisited,” in I. Hacking (ed.) *Exercises in Analysis*, Cambridge: Cambridge University Press.
- Cohen, G.A. (2005) “Facts and Principles,” *Philosophy & Public Affairs* 31: 211-45.
- Copp, D. (2003) “Why Naturalism?,” *Ethical Theory and Moral Practice* 6: 179-200.
- Cuneo, T. (2007a) *The Normative Web: An Argument for Moral Realism*, Oxford: Oxford University Press.
- (2007b) “Recent Faces of Moral Nonnaturalism,” *Philosophy Compass* 2: 850-79.
- and Shafer-Landau, R. (2014) “The Moral Fixed Points: New Directions for Moral Nonnaturalism,” *Philosophical Studies* 171: 399-443.
- Dreier, J. (1992) “The Supervenience Argument Against Moral Realism,” *Southern Journal of Philosophy* 30: 13-38.
- (2015) “Explaining the Quasi-Real,” *Oxford Studies in Metaethics* 10: 273-97.
- (MS) “Is There a Supervenience Problem for Robust Moral Realism?,” unpublished manuscript.
- Enoch, D. (2011) *Taking Morality Seriously: A Defense of Robust Realism*, Oxford: Oxford University Press.
- Fine, K. (1994) “Essence and Modality,” *Philosophical Perspectives* 8: 1-16.
- (2002) “Varieties of Necessity,” in T. Gendler and J. Hawthorne (eds.) *Conceivability and Possibility*, Oxford: Oxford University Press.
- FitzPatrick, W.J. (2008) “Robust Ethical Realism, Non-Naturalism, and Normativity,” *Oxford Studies in Metaethics* 3: 159-205.

- Hampton, J. (1998) *The Authority of Reason*, Cambridge: Cambridge University Press.
- Hare, R.M. (1952) *The Language of Morals*, Oxford: Clarendon Press.
- (1981) *Moral Thinking*, Oxford: Clarendon Press.
- Hills, A. (2009) “Supervenience and Moral Realism,” in A. Hieke and H. Leitgeb (eds.), *Reduction - Abstraction - Analysis: Proceedings of the 31th International Ludwig Wittgenstein-Symposium*, Heusenstamm:ontos verlag.
- Jackson, F. (1998) *From Metaphysics to Ethics*, Oxford: Oxford University Press.
- Kim, J. (1984) “Concepts of Supervenience,” *Philosophy and Phenomenological Research* 45: 153-76.
- Klagge, J. C. (1988) “Supervenience: Ontological and Ascriptive,” *Australasian Journal of Philosophy* 66: 461-70.
- Korsgaard, C. (1996) *The Sources of Normativity*, Cambridge: Cambridge University Press.
- Kramer, M. H. (2009) *Moral Realism as a Moral Doctrine*, Oxford: Wiley-Blackwell.
- Leary, S. (forthcoming) “Non-Naturalism and Normative Necessities,” provisionally forthcoming in *Oxford Studies in Metaethics*.
- Lewis, D. (1983) “New Work for a Theory of Universals,” *Australasian Journal of Philosophy* 61: 343-77.
- McLaughlin, B. and Bennett, K. (2011) “Supervenience,” E. N. Zalta (ed.) *The Stanford Encyclopedia of Philosophy (Winter 2011 Edition)*, URL = <http://plato.stanford.edu/archives/win2011/entries/supervenience/>
- McPherson, T. (2012) “Ethical Non-naturalism and the Metaphysics of Supervenience,” *Oxford Studies in Metaethics* 7: 205-34.
- (2015) “Supervenience in Ethics,” E. N. Zalta (ed.) *The Stanford Encyclopedia of Philosophy (Winter 2015 Edition)*, URL = <http://plato.stanford.edu/archives/win2015/entries/supervenience-ethics/>
- Olson, J. (2014) *Moral Error Theory: History, Critique, Defence*, Oxford: Oxford University Press.
- Ridge, M. (2007) “Anti-Reductionism and Supervenience,” *Journal of Moral Philosophy* 4: 330-48.

- Rosen, G. (MS) "What is Normative Necessity?," unpublished manuscript.
- Scanlon, T.M. (2014) *Being Realistic about Reasons*, Oxford: Oxford University Press.
- Schroeder, M. (2014) "The Price of Supervenience," in *Explaining the Reasons We Share*, Oxford: Oxford University Press.
- (2015) "Being Realistic about Reasons by T.M. Scanlon," *Australasian Journal of Philosophy* 93: 195-98.
- Schmitt, J. and Schroeder, M. (2011) "Supervenience Arguments under Relaxed Assumptions," *Philosophical Studies* 155: 133-60.
- Shafer-Landau, R. (2003) *Moral Realism: A Defence*, Oxford: Oxford University Press.
- Skarsaune, K.O. (2015) "How to Be a Moral Platonist," *Oxford Studies in Metaethics* 10: 245-72.
- Smith, M. (2000) "Does the Evaluative Supervene on the Natural?," in *Ethics and the A Priori*, Cambridge: Cambridge University Press.
- Stratton-Lake, P and Hooker, B. (2006) "Scanlon versus Moore on Goodness," in T. Horgan and M. Timmons (eds.) *Metaethics after Moore*, Oxford: Oxford University Press.
- Sturgeon, N.L. (1999) "The Moral Problem by Michael Smith," *Philosophical Review* 108: 94-97.
- (2003) "Moore on Ethical Naturalism," *Ethics* 113: 528-56.
- (2009) "Doubts about the Supervenience of the Evaluative," *Oxford Studies in Metaethics* 4: 53-90.
- Wedgwood, R. (1999) "The Price of Non-Reductive Moral Realism," *Ethical Theory and Moral Practice* 2: 199-215.
- (2007) *The Nature of Normativity*, Oxford: Oxford University Press.
- Wielenberg, E.J. (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: 595-637.