The Supervenience Objection to Moral Non-Naturalism*

Pekka Väyrynen

University of Leeds

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 actions or their circumstances differ in no other way, such as with respect to 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 are immensely plausible. Each is a way of saying that some normative feature (moral permissibility, moral goodness, reasons for action) is supervenient. So it is immensely plausible that moral features are supervenient. (The same is plausibly true of other normative and evaluative notions as well, such as aesthetic and prudential values.) There are several different relations that go by the name of supervenience, but they all

*Revised draft, 31 March 2015. Please don’t quote without permission. Thanks to Tristram McPherson and David Plunkett for comments on an earlier draft.
share this core idea: things cannot differ in one respect without differing in some other respect. (An excellent overview of supervenience is McLaughlin and Bennett 2011.) 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 rival theories. The supervenience objection to non-naturalist moral realism says that this is the predicament of non-naturalism.

Non-naturalism is a form of moral realism – a family of views according to which there are metaphysically robust objective moral truths. (If a proposition p is objectively true, in the relevant sense, then p is true independently of our attitudes, beliefs, and theories concerning whether p. According to realists, the notion of truth in play 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 the non-naturalist is that these moral properties are sui generis. (Contemporary non-naturalists include Hampton 1998, Shafer-Landau 2003, Cuneo 2007, FitzPatrick 2008, Enoch 2011, Wielenberg 2014, and Cuneo and Shafer-Landau 2014. Some of these authors characterize non-naturalism as “robust” moral realism. Many non-naturalists think that normative properties in general form a sui generis class, of which the moral is a part.)

To characterize normative properties as sui generis is to represent them as discontinuous with both natural and supernatural properties – that is, as neither identical with, reducible to, nor otherwise explicable in terms of natural or supernatural properties. Natural properties are (i) such that 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).
This characterization allows us to count the properties, kinds, and causal systems studied by the natural sciences, psychology, and at least some social sciences as natural, but without ruling out properties to which we have reliable but non-scientific empirical access. Non-naturalists deny that moral properties fall into any such class or the class of supernatural properties (such as the properties of God’s will). This is captured by defining non-naturalist moral realism as the view 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 objection says that non-naturalists cannot explain this connection without 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 objection more carefully and then argue that there isn’t yet a fully effective non-naturalist reply to the objection.

### The Supervenience Objection

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 identify a supervenience claim that isn’t question-begging in the context of the supervenience objection to non-naturalism, and then explain the objection.

If things cannot differ morally without differing in some other 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 does it make whether this duplicate of St. Francis's is actual or merely possible? None, it seems. Accordingly, the supervenience objection to non-naturalism is best understood in terms of strong supervenience, which constrains 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. (Supervenience is non-symmetric. Everything supervenes on itself, and the impossibility of α-differences without β-differences is compatible with the impossibility of β-differences without α-differences.)

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

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

This formula requires both explanation in ordinary language and interpretations of the schematic variables and the necessity operators (marked by ‘\(\Box\)’). I’ll take ‘\(\alpha\)’ and ‘\(\beta\)’ to pick out classes of properties. We’ll be interested in the case where \(\alpha\) is the class of moral properties; so ‘F’ stands for some specific moral property in that class. (Of \(\beta\) more below.) And ‘x’ and ‘y’ pick out individuals like persons or actions. So (SS) isn’t a global supervenience claim to the effect that the world couldn’t have been different in its \(\alpha\)-properties without a difference in its \(\beta\)-properties, but an individual supervenience claim. (For complications regarding strong global supervenience, see McLaughlin and Bennett 2011: §4.4.) It is worth emphasizing that (SS) states an ontological connection between properties, not an ascriptive connection between types of judgments (Klagge 1988). The aim of metaethical expressivists has often been to explain why normative judgments are such as to be governed by a supervenience constraint without an ontological commitment to moral properties (Hare 1952; Blackburn 1985). (Complications that arise in the context of “quasi-realist” expressivism are usefully explored in Dreier 2015.)
The interpretation of the necessity operators is controversial. The innermost necessity is typically taken as metaphysical: whenever something has a moral property, it has a property that metaphysically necessitates the moral one. (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.)

So far we have (SS) as saying the following in ordinary English: as a matter of conceptual/metaphysical necessity (the outermost ‘□’), when something has a moral property, it has some (possibly very 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. So 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). But there is a non-question-begging 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.

This interpretation has two virtues. First, SUPERVENIENCE is suitably neutral between various first-order moral theories because 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. Second, **SUPERVENIENCE** is suitably common ground for the purposes of the supervenience objection. 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. 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 kind of impossibility of a moral difference without a base difference which **SUPERVENIENCE** states is plausibly not brute, so it requires explanation. The version of the supervenience objection to non-naturalism which I’ll discuss is generated by conjoining **SUPERVENIENCE** with a claim concerning what non-naturalism says about **SUPERVENIENCE** and a plausible methodological assumption. (My formulation of the objection will largely follow McPherson 2012: 217-9.)

The methodological assumption in question concerns necessary connections between discontinuous properties. If properties F and G are discontinuous, one might worry that any necessary connection between them, including the sort we find in **SUPERVENIENCE**, isn’t going to be intelligible. That worry goes away if such a connection has a compelling explanation. But if the necessary connection is brute, and so has no explanation, that would seem to be a cost to any view that posits it. This methodological principle may be stated as follows:
MODEST HUMEAN Commitment to brute necessary connections between discontinuous properties counts significantly against a view. (McPherson 2012: 217.)

MODEST HUMEAN treats brute necessary connections as a real cost to any view that posits such connections but doesn’t automatically rule out such views. It allows that overall such a view might have merits that outweigh the cost. (McPherson 2012: 218.) In these respects MODEST HUMEAN is weaker than a more familiar (and controversial) principle known as “Hume’s Dictum” which rules out any metaphysically necessary connections (brute or otherwise) between distinct entities (Lewis 1983: 366; see also Wilson 2010).

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 in the sense at issue in MODEST HUMEAN. But if moral properties are discontinuous with base properties, isn’t it surprising that things cannot differ morally without differing in base respects? Why should that be impossible? So 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 objection to moral non-naturalism differs from others that wear the same label. Perhaps the most discussed version of the objection 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;
This is an overreach: supervenience doesn’t raise problems for naturalist moral realism (Dreier 1992; Sturgeon 2009). The supervenience objection 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 objection to moral non-naturalism. I’ll first consider the prospects of rejecting BRUTE CONNECTION and then discuss whether non-naturalists might avoid the supervenience objection 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.)

**Rejecting BRUTE CONNECTION?**

Among the three premises that generate the supervenience objection to non-naturalism, BRUTE CONNECTION seems to be 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 explanation 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, the flouting of which manifests a conceptual deficiency. The Conceptual Strategy is thus to explain SUPERVENIENCE by the nature of moral concepts. 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 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. If (as some philosophers have done) you regard it as a conceptual necessity that magnets attract iron, you’re not done. You still owe a substantive explanation of why or how magnets do this. Realists about magnets and iron cannot stop with saying that the *concepts* of magnets and iron guarantee attraction between them. That must be up to the *world* to decide, not our representations of it. 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.) It seems plausible that SUPERVENIENCE is 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 it is analytic. And since SUPERVENIENCE is a connection between between classes of *properties*, presumably it holds (if it does) irrespective of whether it is reflected in moral *concepts*. (That is, it would 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 \textit{Supervenience} contain two necessity operators: \textit{necessarily}, if any two possible individuals are alike in every base respect, then they \textit{must} be alike in every moral respect. Now, when proponents of the Conceptual Strategy say that \textit{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. The innermost necessity operator is typically interpreted as metaphysical: when something has a moral property $M$, there is some or other (possibly complex) base property that attaches to $M$ with metaphysical necessity. So the claim which the Conceptual Strategy represents as a conceptual truth is a truth about a metaphysical necessity. But what the supervenience objection to non-naturalism represents as needing explanation is the metaphysical necessity by which the base property necessitates $M$. The Conceptual Strategy thus explains the wrong necessity. The puzzle for non-naturalism remains even if \textit{Supervenience} is a conceptual truth: if moral properties are \textit{sui generis}, how can the necessary metaphysical connection between non-moral base properties and moral properties (the innermost necessity operator) be anything but brute?

A natural alternative to the Conceptual Strategy is to seek a metaphysical explanation of \textit{Supervenience} – a Metaphysical Strategy. Think of the challenge this way: how could there be properties which are \textit{sui generis} and yet supervene in the way \textit{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 it must be explicable, because explaining one necessary connection by relying on another brute necessary connection 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 composed 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 composed of different non-moral facts, then generosity isn't identical with or reducible to 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 a moral property is wholly constituted by some concatenation of non-moral properties, then its nature is being understood in terms of non-moral properties – and therefore as not discontinuous with them. Second, explanations that rely on considerations that are murkier than what they aim to explain are no good. We understand constitution in the case of material objects, such as what it is for a ship to be constituted by its parts. What it is for one fact to be constituted by others is much less clear. If a fact is an instantiation of a property by an object (at a time), the kind of structure that constitution would require facts to have can only be found in their property component. One problem here is that the claim that one and the same *property*, F, is constituted by G, H, and I on one instantiation but by J, K, and M on another is difficult to make sense of. It would be much more straightforward to say that one object is F in virtue of being G, H, and I and another object is F in virtue of instantiating J, K, and M. This requires no additional claim about F itself. (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 fare no better. Suppose we say instead that moral properties are *exclusively realized* by non-moral properties, on analogy with how mental properties like being in pain can be realized by different physical substrata (Shafer-Landau 2003: 77). This
analogy fails. Different realizers of a multiple realizable property exclude one another; for instance, if a subject’s beliefs are realized by brain states, they cannot be realized also by green slime states. But super-virtuous people who instantiate multiple bases of moral goodness are possible. 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 exhausted by these options. One might, for instance, think that it is part of 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 flag this kind of explanations of supervenience for discussion elsewhere. (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. The foray into modal logic required to assess this account of supervenience would take us too far afield; but see Schmitt and Schroeder 2011.)

The last response to brute connection which I’ll consider is the Tolerable 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 this strategy is T. M. Scanlon’s explanation of normative supervenience. (Scanlon thinks that normative facts exist in an ontologically non-committing sense, so he isn’t a “robust” non-naturalist realist.) 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 it 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 <p, x, c, a>. Tuples that stand in R thus have their non-normative features necessarily. So when R holds of <p, x, c, a>, 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 also offers to non-naturalists an explanation of supervenience based on a conception of basic normative truths on which such truths are always necessary.)

One worry about the Tolerable Substitute Strategy is that it 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 <p, x, c, a> is just to
say that there are some necessary relationships that hold among wholly dis-
tinct entities“ (Schroeder 2015: 197). Considerations that motivate MODEST
HUMEAN might then be thought to apply with similar force to the necessary
relationships that are crucial to Scanlon’s account.

This worry can be recast more effectively as a general worry about treat-
ing normative truths as brute which needn’t appeal to MODEST HUMEAN.
The Tolerable Substitute Strategy aims to make it less puzzling how two dis-
continuous 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, what-
ever they are, in terms of a set of particular pure normative truths which
delivers a corresponding 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 <p, x, c, a> instead of some other tuple
<q, y, d, b>? The distribution of the R relationship over facts, agents, cir-
cumstances, and actions shouldn’t be arbitrary. The same question extends
to the necessity of these relationships: when <p, x, c, a> 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.
(Contingency is at least the default status for truths. 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. (This sort of explanatory demand is discussed in Schroeder 2007:
69-71 and Väyrynen 2011: 185-92.) Those who find the demand legitimate
are bound to regard Scanlon’s explanation of SUPERVENIENCE as leaving too
many significant normative truths unexplained.
Rejecting \textit{SUPERVERIENCE}? 

A different way to defend moral non-naturalism against the supervenience objection is to reject \textit{SUPERVERIENCE}. This strategy might seem quixotic since \textit{SUPERVERIENCE} seems to be supported by highly compelling intuitions, such as those with which we began. But if these intuitions could be captured otherwise, there would be room to reject \textit{SUPERVERIENCE}. Call this the Moral Contingency Strategy. (The label is due to Dreier MS. If \textit{SUPERVERIENCE} is false when the innermost necessity is read as metaphysical, then even the most basic moral principles are metaphysically contingent.) I’ll argue that this strategy gives the non-naturalist no distinctive explanatory advantage, and therefore doesn’t warrant rejecting \textit{SUPERVERIENCE}.

The most developed form of the Moral Contingency Strategy is due to Gideon Rosen (MS). Rosen argues that \textit{SUPERVERIENCE} and non-naturalism form an inconsistent triad with “essentialism” about metaphysical modality, and that out of these three claims we should reject \textit{SUPERVERIENCE}. Strong supervenience entails that for each moral property \(M\) there is a condition consisting of a set of base properties, \(\Phi\), which is equivalent to \(M\) as a matter of (metaphysical) necessity. At minimum \(\Phi\) 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 \(\Phi\) 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 \(\Phi\) such that for some item \(x\), the equivalence between \(M\) and \(\Phi\) 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 a distinctive commitment of moral non-naturalism is 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 is it SUPERVENIENCE that we should reject? Why not put essentialism in the reject pile instead? Essentialism requires very substantial metaphysical commitments regarding the essences of things and casts metaphysical necessity more narrowly than many philosophers think. But if you are happy to grant essentialism, at least for argument’s sake, now ask yourself: in virtue of what item’s nature might the necessary equivalence between (say) moral rightness and \( \Phi \) hold? It won’t be anything about \( \Phi \). For if the properties in \( \Phi \) are wholly non-moral, then their natures are silent about morality. Nor will it be anything about moral rightness. For rightness won’t have heard of many of the non-normative properties and relations in \( \Phi \), 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 \( \Phi \). But given essentialism, the metaphysically necessary equivalence between \( \Phi \) and moral rightness fails if it is logically consistent with all the essential truths that something is wrong but satisfies \( \Phi \). And if the necessary equivalence fails, SUPERVENIENCE is false.
The point can also be brought out using 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, then surely absolutely anyone exactly alike him in their non-moral properties would 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 2005 introduces a similar notion of the fact-independence of moral principles.) 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; for discussion, see Cuneo 2007b: 863-71). (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 also an instance of normative necessity; Scanlon 2014: 41 n. 40.) 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 still 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.

This argument offers powerful reasons to regard fact-independence as an important feature of moral principles and to regard supervenience, understood as constraining all metaphysically possible individuals, as controversial in its own right and not required to make sense of the practice of moral
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 domain-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.) I’ll close with three critical observations about the Moral Contingency Strategy.

First, the Moral Contingency Strategy doesn’t help non-naturalism in particular because its characterization of moral naturalism is too narrow. The distinctive commitment of naturalism is that moral facts and properties belong to the category of natural facts and properties. It doesn’t follow that a property can count as natural only if it can also be ascribed or represented in some purely non-normative terms (Sturgeon 2003: 536-40). So moral naturalism as such entails no commitment to a necessary equivalence between each moral property and a condition specifiable in wholly non-normative terms. Some forms of naturalism require only a trivial sort of supervenience: moral properties cannot change without a change in some natural properties, namely in themselves.

Second, the Moral Contingency Strategy leaves something morally important unexplained. If act utilitarianism is the basic principle of morality in some possible worlds, but not ours, why is that? If unalloyed racism, sexism, and violations of others’ bodily integrity are moral obligations in some metaphysically possible worlds, how come isn’t ours one of them? Rosen’s argument implies that these questions have no answer; the metaphysical contingency of the basic moral principles governing a given world is brute. But insofar as contingency is the default status of truths, shouldn’t truths that are normatively necessary also generally require explanation? If this demand for explanation is legitimate, then failure to explain normative necessities would still be a strike against a theory that posits them. Should no such explanation be forthcoming from the Moral Contingency Strategy, the supervenience objection to non-naturalism might retain some of its bite even if we rejected strong supervenience claims that constrain moral differ-
ences across all metaphysically possible worlds. This worry raises some very general issues. (In metaphysics and philosophy of science, it is widely accepted that laws of nature support counterfactuals but controversial whether these laws are grounded in essences or otherwise metaphysically necessary. Rosen’s metaethical position is analogous to thinking that laws of nature are explained by some metaphysically contingent facts. The question to explore will be what metaphysically contingent facts are supposed to explain the important modally restricted claims which morality supports – such as that if murder is wrong, then murder would have been wrong no matter how the non-moral facts had been – and whether they do so.) But 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 Rosen’s argument unsatisfactory. (One explanation of such normative questions can be found in Korsgaard 1996.)

Third, the Moral Contingency Strategy entails too much moral luck. (The argument to follow is due to Dreier MS.) Given moral contingency, it is metaphysically possible that the world could have been just as it is in wholly non-moral respects but any, or even all, of 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 accept the same moral principles we actually accept, think of ourselves as morally imperfect but fundamentally decent persons, and so on; in counter-moral possible worlds we would just be very badly mistaken to think these things. 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 grotesquely evil. Yet that is our situation, according to the Moral Contingency Strategy. So we should reject the Moral Contingency Strategy.

In conclusion, the Moral Contingency Strategy is intriguing and deserves further discussion both in its own right and as a response to the supervenience objection 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 objection to non-naturalist moral realism and assess some of the most promising responses to it. I focused on two main strategies: rejecting the relevant supervenience thesis itself (SUPERVENIENCE) and rejecting the claim that non-naturalists must take moral supervenience to involve brute necessary connections between discontinuous properties (BRUTE CONNECTION). Although some of these responses involve fascinating philosophical ideas, I nonetheless argued that some of these responses are ineffective and that even the most promising attempts suffer from some explanatory shortcomings. The demands for explanation on which these shortcomings are premised aren’t themselves entirely uncontroversial, however. The supervenience objection to non-naturalism is far from a closed chapter in metaethics.
Related Topics

TBD

References

— (MS) “Is There a Supervenience Problem for Robust Moral Realism?,” unpublished manuscript.


Rosen, G. (MS) “What is Normative Necessity?,” unpublished manuscript.


