

4

Bruteness and Supervenience

Mind vs. Morality

Joseph Levine

In this chapter I will first present a framework, one I have defended elsewhere (Levine 2001), for understanding the notion of bruteness, its relation to modality, and the way this framework applies to the mind–body problem. Second, I will explore the challenge that the combination of views known as “robust realism” or “non-naturalism” about the ethical and the supervenience of the ethical on the non-ethical, both views I am inclined to endorse, poses to this framework. In response I will speculate about how one might reasonably meet the challenge.

1.

As far as I can tell there is no standard meaning for the term “brute” as in “brute facts”. For the purposes of this chapter I want to use “brute” to indicate an epistemological feature, specifically having to do with explanation, reserving “basic” for the corresponding metaphysical feature. In this section I will explain how I understand these two features and also present and defend certain fundamental principles I hold regarding their nature and relationship.

Let me begin with the metaphysical side. First of all, to what does the adjective “basic” apply? One might choose statements, propositions, facts, states of affairs, etc. I opt for fact-like entities I call “situations”, which, when they obtain, are facts. A situation, then, is a structure of objects, properties, and relations. For instance, the situation of my computer resting on my desk is a situation involving my computer, my desk, and the relation of resting on.

So, some situations are possible, some are actual (and thus, obviously, possible), some necessary, and some impossible. My computer being on my desk at the present time is both actual and possible, though at other times it is merely possible. My computer being self-identical is necessary, and my computer both being on my desk and not on my desk at the same time is impossible. Using the possible worlds framework, merely possible

situations obtain in some possible world, but not the actual world, necessary situations obtain in every possible world, and impossible situations obtain in no possible world.

The kind of modality at issue here is metaphysical. The space of worlds that are possible in the broadest sense is the space of metaphysically possible worlds. I understand this to be the core meaning of possible, with other forms of possibility, like nomological possibility, involving a restriction on the space of metaphysically possible worlds. I am inclined to use “metaphysical possibility” and “logical possibility” interchangeably, since, as will become clear presently, I believe that logic is the primary determinant of possibility and necessity. However, since others use these terms differently, and I want to avoid any confusion, I will stick to “metaphysical possibility” (and necessity); and wherever I use the terms “possible” and “necessary”, unless otherwise explicitly indicated, it is metaphysical possibility and necessity that I intend.

So what makes a situation possible or not? I don't think it helpful here to advert to possible worlds and say that what makes a situation possible is that it obtains in at least one possible world. Rather, if anything, the determination relation here is the reverse. The space of possible worlds is what it is because of the situations that are possible and impossible. It is not as if there is an independently given space of possible worlds and we, as travelers through modal space, can come to discover, as it were by rational observation, what happens to inhabit this or that world. Rather, the framework of possible worlds is just a heuristic device for representing possible and impossible situations.

So what does determine, or ground the possibility or impossibility of a situation? Here is where logic comes into the picture. What would be straightforward is to say that a situation is possible just in case it isn't ruled out as logically inconsistent. However, consistency is a property that applies to statements, or representations of situations, and so first we need to say something about how representations of situations fit into the picture. Situations, as I have defined them, are structures of objects instantiating properties and standing in relations. But situations have descriptions in language, both public and mental. I will assume for now that our cognitive activity—thoughts, beliefs, inferences, etc.—is mediated through a system of mental representation that is syntactically structured and semantically interpreted; that is, I endorse a Language of Thought hypothesis. For each situation (at least finite ones) there correspond representations that describe it, with singular terms referring to the objects constitutive of the situation, predicates representing the properties and relations involved, and the requisite syntactic structure to capture the structure of the situation. I say representations, plural, because representations stand to objects, properties, and situations in a many-one relation.

Logic, as mentioned above, in the first instance applies to representations. A representation can be logically valid, consistent, or inconsistent, and some representations follow from others and some are inconsistent with each other. Consider the representation “JL's computer is both resting on his desk and not resting on his desk.” Once we have the intended interpretation for the primitive terms in the representation it is apparent that the situation represented by this sentence is impossible. Is the converse true? Does it follow from a situation's being impossible that a representation of it will

be inconsistent? No. After all, there are many different representations of any single situation, and even if the situation itself is impossible only a subset of its representations will manifest its impossibility by being logically inconsistent. So, for instance, the very same impossible situation represented by “Hesperus is not identical to Hesperus” can be represented by “Hesperus is not identical to Phosphorus.” Since both names, “Hesperus” and “Phosphorus” refer to the same object, the two statements describe the very same situation. But only one of the two statements is inconsistent, and so only it reveals the situation’s impossibility. This is where the metaphysics and epistemology of modality diverge somewhat; this divergence is a consequence of the multiplicity of representations picking out individual objects, properties, and situations.

To complete the picture of metaphysical modality and its epistemology, I need to introduce one more notion: what it is for a situation to be conceptually possible (or epistemically possible, or conceivable—I will use these terms interchangeably). A situation *S* is conceptually possible relative to representation *R* just in case *R* represents *S* (which, again, means that the singular terms in *R* pick out the objects in *S*, the predicates in *R* represent the properties and relations in *S*, and the syntactic structure of *R* represents—according to the compositional semantics of the representational system containing *R*—the structure of *S*) and *R* is logically (and conceptually)¹ consistent. Thus where metaphysical possibility (or plain old possibility) is a feature of situations, conceptual possibility is a matter of the relation between representations and situations. In other words, the very same situation *S* can be conceptually possible relative to *R* but conceptually impossible relative to *R*’.

What do we say about a situation that is conceptually possible relative to one representation but conceptually impossible relative to another? What does this tell us about its metaphysical status? Well, if a situation can be represented by any logically inconsistent representation I say it is metaphysically impossible. Above I said that logic is the determinant of metaphysical modality, and yet that the latter applies in the first instance to situations and the former in the first instance to representations. So how, one might ask, could the former be the ground, or determinant of the latter? The answer is to be found in the connecting principle just mentioned: I say a situation is possible just in case there does not exist a representation of it that is logically inconsistent. So, for instance, the situation of John’s being a married bachelor is impossible because that very same situation has the following description: “John is both married and unmarried.”

On my view, then, metaphysical possibility and necessity is grounded in logical possibility and necessity. In this respect I adhere to the doctrine Chalmers (1999) calls

¹ I added “conceptually” in parentheses to allow for a priori knowable analytic truths, if there are any. Any situation that is ruled out on conceptual grounds could also be ruled out on logical grounds, assuming the relevant analytic truth supports an identity claim. So if bachelors are unmarried men and water is H₂O, then both (1) “John is a married bachelor” and (2) “This water contains no hydrogen” represent impossible situations, since they represent the very same situations as (1’) “John is a married unmarried man” and (2’) “This group of H₂O molecules contains no hydrogen.” However, if “Bachelors are unmarried men” is analytic, knowable a priori, then I wouldn’t want to call the situation represented by (1) conceivable, even relative to that representation, whereas we standardly do call the situation represented by (2) conceivable.

“modal rationalism”. This is the idea that modality fundamentally describes the limits of rational conception, not an independent reality. This is why I said the space of possible worlds is not independently specifiable, as, say, the worlds in a galaxy or the universe as a whole would be. Rather, talk of possible worlds is determined by what is (logically, or conceptually) consistently conceivable. However, due to the many–one relation holding between representations and situations—as famously emphasized by Frege—one cannot immediately read off a description of a situation that it is possible. So long as there exists at least one representation of it—even one we aren’t aware of—under which it is logically inconsistent, the situation isn’t possible, no matter how much it seems so. This, to my mind, is the source of the division between metaphysical and epistemic modality; it is a consequence of the fact that there can be distinct representations of the same thing without our knowing it.²

After this detour through the metaphysics and epistemology of modality, let’s return to fundamentality (or being basic—I will use the terms “fundamental” and “basic” interchangeably). I call a situation basic if it is not realized or constituted by another situation (or set of situations). So what’s doing the work here is the notion of realization (or constitution). It is extremely difficult to provide a non-circular characterization of the realization relation, so I will merely provide examples that have been used for decades to illustrate the notion in philosophy of mind.

Functionalism is the doctrine that mental states are functional states; that is, states that are causal roles of sorts. A causal role is realized by a set of objects and properties when their causal powers meet a canonical specification of the causal role. So, to use one of the standard examples, a speedometer is realized by the particular mechanical and electronic parts that compose it because their being constructed as they are results in certain of the speedometer’s states (the numbers displayed) being causally correlated with the speed of the vehicle. Similarly, the electronic components and properties of a computer realize a program because they determine a sequence of states that meet the specifications of the program.

Probably most of the situations we encounter in daily life and represent in our cognitive activity are realized by more basic situations. So, for instance, when I judge a piece of fruit ripe, the situation so described is realized by a very complex collection of chemical situations. Deleting a word from my document is realized by a similarly complex collection of electronic situations taking place in my computer. My being a professor at UMass Amherst is realized by a dizzyingly complex collection of situations involving people, their intentions, their movements, and the like.

² There are two kinds of cases that might seem counterexamples to my “connecting principle”. First, if “the cause of the broken window” refers to the impact of the stone, then “the cause of the broken window didn’t break the window” might seem to be a logically inconsistent description of a possible situation. Second, “Aristotle was not human” seems to represent an impossible situation, yet it’s not obvious there is any logically inconsistent representation of that situation. See Levine (2001) where I address these challenges.

What's crucial about realization or constitution, as I understand it, is captured by the phrase "nothing over and above". When a situation A is realized by situation B (or set of situations B1...Bn), it is appropriate to say that A is nothing over and above B. Sometimes the metaphor of God's creation is used to elaborate on this idea of one situation being nothing over and above another. So, it is said, if A is nothing over and above B, then when God created B, or put B into the world, there was nothing more for Him to do in order to also bring A into the world. When God created all the right chemicals and the laws governing them, there was nothing further He had to do to make fruit ripe.

The connection, then, between fundamentality (or basicness, if that's a word) and the "nothing over and above" locution is this: Basic situations are those that are something over and above other situations. If one situation is nothing over and above another, then it is because the latter realizes the former, which makes the former not basic.

Finally, we need to introduce supervenience. To say that properties (or situations) of kind A supervene on properties of kind B is to say that there can be no difference in situations, or worlds, with respect to the A properties without there being a difference in the B properties. Put conversely, the distribution of B properties (or situations) metaphysically necessitates the distribution of A properties.

Clearly, when the A facts are realized by the B facts, when they are nothing over and above the B facts, then they also supervene on the B facts. However, given the definitions of "realization" and "supervenience" the converse doesn't follow: it is not the case that if the A facts supervene on the B facts then it follows that the A facts are realized by the B facts. Whether one set of facts can supervene on, and thus be metaphysically necessitated by, another set while still being something over and above them will be the focus of my discussion about ethical supervenience below.

In philosophy of mind the dispute between materialists and anti-materialists can be put this way: For materialists, mental situations (or properties) are not basic; they are always realized by physical (or non-mental) situations. If pain is a functional state realized by c-fiber firing in organisms whose overall physical constitution meets the appropriate conditions, then pain is nothing over and above c-fiber firing in the context of these other physical conditions. Anti-materialists claim that pain is not realized by this configuration of physical states because one needs to add certain phenomenal properties that are not realized physically. That is to say, phenomenal pain is a basic property.

2.

Now let's turn to the notion of bruteness. As I said at the beginning of the chapter, I consider bruteness to be an epistemological notion. To call a fact, situation, or whatever "brute", as I understand it, is to say that it has no explanation. Obviously we don't always know whether there really is an explanation and we just haven't hit on it yet. Sometimes we expect that though we don't know the explanation yet, there is bound to be one and we have some idea what form it will take; it's just that the details are still not discernible.

But sometimes we have no idea what would explain the fact in question and treat it as brute, inexplicable.

The important philosophical question regarding bruteness, as I see it, is determining when it is acceptable to posit brute facts and when it isn't. So, some philosophical (or scientific) theories have as consequences that certain facts turn out inexplicable, or brute. Sometimes this is not considered a serious problem, and we console ourselves by saying "Well, explanation has to stop somewhere, why not here?" But other times the fact that a theory entails that certain facts are brute is considered, if not a knock-down objection to the theory, at least a very high cost of holding it.

Let me describe an example of each case. Consider the magnitude of the gravitational constant. At least on some physical theories this is a brute fact. Even if it turns out to be explicable in terms of more basic physical facts and laws, at some point we know that explanation will come to an end. When we encounter the most basic, fundamental laws of our universe, we are inclined to say they are the way they are just because that's how they are. End of story.

Another, more controversial example of acceptable bruteness comes from my own criticism (Levine 2016) of an argument of Robert Adams (1987). Adams takes the explanatory gap between physical and functional facts on the one hand and phenomenal facts on the other to be a strong consideration in favor of theism. His argument, briefly, is this. Given that there is no explanation we can imagine for why certain physical/functional facts should always be correlated with certain phenomenal experiences (such as what it's like to see red or feel pain), positing God as a creator of the world and its laws removes the unacceptable bruteness of the correlation. Even though we don't have access to God's reasons, still we know that when certain physical/functional configurations give rise to experiences of phenomenal red, it was by God's design. We would have confidence that there is an explanation for why just this phenomenal property and not some other is tied to this particular physical situation.

However, I argued as follows. The explanatory gap does indeed lend strong support to some form of dualism. However, once one posits that there are psycho-physical laws that relate physical situations to phenomenal ones, and that these laws are basic, not realized in purely non-psychological laws (as would be the case if materialism were true), I argue that concerns about the bruteness of said laws lose their force. After all, if physical facts and laws that are basic do not require an explanation, why should basic laws that relate physical situations to mental ones require it? In other words, basicness, or fundamentality, is, as we might put it, a legitimate excuse for bruteness. No need to bring God in to save the day.

So one lesson we learn from this case, if you buy my argument, is that being metaphysically basic is an acceptable explanation for being explanatorily brute. Most of the individual facts and regularities we encounter in the empirical world are presumably not basic, so therefore we expect them to be explicable, not brute. But when we come to the most basic facts and laws that characterize our world, we assume that there is no explanation—it's just what our world is like. It could have been different, but it isn't.

This last point, that it could have been different, brings up an important dimension to explanation. It seems to me that explanation is largely a matter of removing possible alternatives. In the limit this would entail showing how the explanandum logically follows from the explanans, though we rarely get such complete explanations. Usually all that's necessary is that the explanans logically exclude (or logically make much less probable) relevant alternatives, the ones that are in the theoretical game, as it were, and not all possible alternatives. But, I claim, if we have a really complete explanation for a situation or law then it should literally follow from the explanans.³

The idea is this. Suppose there is some phenomenon we want explained. Why did the baseball break the window? If we say something general like "because it's solid and glass is fragile so when the ball hit the glass it broke" this may be fine as far as it goes, but the explanation is open to an objection. Well, one might say, not every time a solid object makes contact with a fragile object does it break. So, for all you've said so far, we still don't know why the glass broke on this occasion, rather than not breaking, as happens on other occasions. If we then add detail about the momentum with which it hit the glass and the glass's threshold for absorbing a shock without breaking, then we've shown why in this particular situation it *had* to break. It isn't until we show how it *had* to happen, given the particular facts and general laws, that we have completely explained this occurrence. Of course, as I said above, as a practical matter we never give truly complete explanations, but the point is that as we consider relevant alternatives we strive to rule them out by showing they couldn't have happened, which explains why they didn't, and the occurrence in question did instead.

So when I say that basic situations and laws are not amenable to explanation, the idea is that these basic facts are what distinguish our world from other possible worlds. In this world, birds fly but humans don't. In other possible worlds it's the reverse. Why do birds fly and not humans? Well, there's a lengthy explanation adverting to physical and biological facts and laws that does explain why, given the basic laws of this world (and its initial conditions, perhaps), birds fly but humans don't. However, since other worlds are possible, and explanation involves showing why these apparent possibilities aren't really possible, we have to expect that the basic facts that identify the actual world as the world it is cannot be explained. To explain them would be tantamount to showing how the actual world is the only possible world; which, given what I've said about the relation between logic and necessity, would mean showing that there is a description of the basic structure of our world that is logically valid.

On the other hand, though dualists, I claim, have no problem considering the correlation between physical facts and phenomenal facts to be brute, materialists do. For dualists, as explained above, this correlation is underwritten by a basic psycho-physical law of nature, and so, as with all basic facts, it is not amenable to explanation. Indeed, the idea of positing this correlation as a nomological necessity is to allow that it may not hold in other possible worlds. Hence, why this world is different from other worlds

³ I defend this thesis more fully in Levine (2001).

is left as a brute fact. That's just what it is for this world to be the world it is; it is for it to have this particular collection of basic facts (and laws), in terms of which everything about this world is explained, but are themselves inexplicable.

However, materialists must claim that mental facts are realized by physical facts, hence there can't be any basic laws or facts that involve the mental. As mentioned above, realization entails supervenience, which is a metaphysically necessary relation. That is, if the mental supervenes on the physical, then there can be no difference in mental state without a difference in physical state, and this holds across possible worlds. Given that the relation between the physical facts and the mental facts isn't confined to this world, it requires explanation, it can't be brute. Thus, if there is indeed an explanatory gap between the physical and the mental it is a problem for materialists.

3.

Having laid out the crucial notions involved here, I want now to develop in more detail their application to the case of mind, which will set us up for the challenge from ethics in the next section. As discussed above, to say that the mental supervenes on the physical is to say that no two situations can differ mentally without differing physically. There are two clarifications that need to be made here: what are included in the relevant situations and what is the nature of the modality at issue.

Regarding the first question, here is the issue. If we take as the different mental situations two beliefs and the physical situations the believers' brain states, then many materialists would argue that there indeed could be differences in beliefs without corresponding differences in brain states. After all, the point of the Putnam Twin Earth example is precisely to insist on this possibility. Oscar can believe that water is wet on Earth while Toscar believes that twater is wet on Twin Earth and yet the brain states of both Oscar and Toscar are the same.⁴

However, if we include in the supervenience base for belief the physical situations that constitute the environments in which the cognitive abilities of the believers in question develop and occur, then it won't turn out in the Oscar–Toscar case that we have a violation of supervenience. It's because Oscar lives on Earth, where H₂O is the local watery stuff, and Toscar lives on Twin Earth, where XYZ is the local watery stuff, that their beliefs differ. So, in general, when considering whether a difference in mental state can occur without a corresponding difference in physical state, we need to include all those aspects of the physical environment that materialists might include, depending on their philosophical theory of mind. Nothing in our discussion depends on how wide a swath of the physical world is supposed to constitute the supervenience base of the mental, so long as it's clearly physical (or non-mental).

The second question has to do with the modality at issue. Traditionally a distinction is made between so-called “weak supervenience” and “strong supervenience”.⁵ Weak

⁴ Of course not everyone accepts externalism. For instance, see Segal (2000).

⁵ For the various versions of supervenience see Kim (1993).

supervenience takes the claim of no difference in the mental without a difference in the physical to apply within possible worlds whereas strong supervenience takes it to apply across possible worlds. So, if we imagine Oscar and Toscar in the same possible world, but on different planets, then according to the doctrine of weak supervenience it can't be that Oscar and Toscar differ in their beliefs without their being a physical difference in their situations; for example, that one is surrounded by H₂O and the other by XYZ. However, if we take them to be located within different possible worlds no such requirement of physical difference exists. On the other hand, the doctrine of strong supervenience maintains that whether Oscar and Toscar are located within the same possible world or not, there had better be some physical difference in their situations if they differ in beliefs (or any other mental state, obviously).

In the literature it's usually assumed that strong supervenience is what is at issue. For one thing, weak supervenience seems compatible with some versions of dualism, especially emergentism. If there are basic laws connecting the physical to the mental, then weak supervenience but not strong supervenience would hold. Given a world, and its laws—including its basic psycho-physical laws—you won't have two creatures who differ mentally without differing in their relevant physical properties. However, when comparing two creatures in two different possible worlds, there could be sameness of physical properties without sameness in mental ones, because the basic psycho-physical laws could differ. At any rate, since it's only strong supervenience that posits a necessitation relation between the physical and the mental, this is the sort of supervenience with which I'll be concerned.

So, assume that the mental strongly supervenes on the physical. This means that if all of the relevant physical properties are fixed, then that (metaphysically) necessitates the mental properties. Now, the natural question to ask is what explains this necessitation relation? There are two choices: either there is an explanation or the fact that the physical necessitates the mental is a brute fact. Let's consider the second option: is it acceptable to claim that this is just a brute fact?

Earlier we claimed that the only acceptable way to explain bruteness is to appeal to basicness, or fundamentality. In our example we claimed that the magnitude of the gravitational constant is a brute fact because it is one of the fundamental features of our world, part of the basic fabric of laws that constitute the world we live in. We don't expect to be able to explain such basic features—indeed, this is part of what is entailed by calling them basic, or fundamental.

Could we just say that the fact that the physical necessitates the mental in the way it does is a basic fact, like the gravitational constant? It just turns out that possible world space is limited in just this way, that no two worlds can differ in their distribution of mental properties without differing in their distribution of physical properties. It's a basic fact about the structure of possible world space.

I, along with a large number of philosophers, find this idea repugnant. While I am quite willing to grant that there are basic, and therefore explanatorily brute, facts and laws that distinguish various possible worlds, the idea that there should be such facts about the space of possible worlds itself strikes me as bizarre, and almost unintelligible.

Some put this point in terms of “plenitude”, the idea that there is nothing arbitrary, or gappy about possible world space.⁶ Others, especially Chalmers, put this in terms of there being an inherent connection between modality and rationality, which is why he calls his position “modal rationalism”.

The basic idea, as I see it, is this. When it comes to actuality, we expect that we need to discover through empirical interaction with the world which situations obtain and which don't, since, as Hume so well puts it in the *Dialogues on Natural Religion* (Hume 1998), before encountering the world, for all we know it could be any consistent way we might imagine it. Once we encounter it, we remove more and more possibilities as we discover more and more. Ultimately we bump up against the actual world's basic structure—one big possibility among many—and say, okay, this just happens to be the way the world is. Not much more you can say about it.

But what would it mean to say that the space of possible worlds “just happens” to be the way it is, limited in various ways that cannot be explained because they are basic? Again, as I understand “just happens to be”, it is a matter of one possibility among many being actual. The gravitational constant could have been any value, but it just happens to be this one. But again, when it comes to what's possible, there's nothing else that includes it from among which it just happens to be selected. What's possible defines the “selection space” as it were, and so can't itself be basic and brute.

So what does define the limits of possibility? Well, as I see it, logic does (and the structure of our concepts, if there are analytic truths). Both sources are a priori and therefore involve no bruteness or arbitrariness, and don't require encounter with the empirical world to determine. If you want to ask, what explains the truth of logic, or underwrites conceptual truth, I'm willing to accept these as brute if you want to call them that. But the bruteness here is of a very different character. Remember, explanation, as discussed above, is a matter of removing alternative possibilities. Why P, when it could have been Q or R? But when we entertain P and see that it *has* to be that way, it couldn't have been otherwise, explanatory questions naturally come to a halt.

So if brute necessities of this sort are unacceptable, and we accept the supervenience of the mental on the physical, we are faced with the need to explain it. What does explain it? Well, basically, the claim that the mental is nothing over and above the physical explains it. Intuitively, if the mental *is* nothing over and above the physical, then how could two creatures differ mentally without differing physically (including their relations to other physical things, as emphasized above)? While this intuitive explanation might be enough for our purposes, it can be reframed in terms of the two bases for necessity just described: conceptual structure and logic.

Now it's precisely the absence of a conceptual or logical connection between descriptions of physical facts and descriptions of mental (particularly, phenomenal) facts that has fueled recent anti-materialist arguments. Chalmers (1996)—whose doctrine of modal rationalism I said earlier I largely endorse—has pressed just such an argument.

⁶ See Lewis (1986).

He claims that if materialism is true, then strong supervenience of the mental on the physical must hold, and so it requires explanation by appeal to either a priori conceptual connections or logic. However, since “zombies”—creatures that are physically just like us but who do not have phenomenal experiences—are coherently conceivable, there must not be the right sort of conceptual or logical relation between the two descriptions. Either we reject the supervenience claim (his option), and therefore reject materialism, or we allow just the sort of “brute necessity” we claimed was unacceptable.

However, as I have argued elsewhere (Levine 2001, 2014), I don’t think the conceivability of zombies does lead to an unacceptable form of brute necessity. Remember above I said that if A necessitates B what’s necessary, to avoid bruteness, is that there be some representation of A and some representation of B under which B follows from A. So, if materialism is true, then presumably there is some physical description of a creature with specific phenomenal experiences in non-phenomenal terms such that that description follows from the one that specifies all of the physical facts regarding that creature. For example, the phenomenal experience of red might be identical to a particular functional state. If so, then the fact that the relevant physical facts necessitate the subject’s having an experience of phenomenal red is explained by the fact that that experience can also be described in functional terms and the latter description is derivable from the underlying physical description. The conceivability of zombies, on this view, is a symptom of the general fact that two representations can pick out the same thing even though there is no a priori connection between them. Of course there is a lot more to be said on the matter, but this is not what I want to pursue here.⁷

4.

Having developed the framework for determining when bruteness is benign and when it’s not, let’s turn now to the challenge from ethics. In the last decade or so there has been a growing literature on the problem that supervenience poses for ethical non-naturalists.⁸ In particular, an argument can be made that, using the framework developed above, the ethical non-naturalist who endorsed supervenience would be committed to an inexplicable brute necessity. I will present the problem and then explore a possible solution.

To simplify matters, let’s assume that there are three ethical properties that apply to actions—right (ought to do), wrong (ought not to do), and permissible (may do or not do). The doctrine of ethical supervenience states that no two actions (across possible worlds) that agree in all of their non-ethical properties can differ in their ethical properties.

⁷ As I argue in Levine (2014), while I don’t accept the Chalmers argument, because I don’t think even “type-B” materialists are guilty of positing brute necessities, I do think there are explanatory considerations that make non-materialism the most plausible hypothesis.

⁸ For a comprehensive review of the literature on this topic, see McPherson (2015).

While not everyone accepts the supervenience of the ethical on the natural, it is endorsed by most meta-ethicists.⁹ The idea is compelling. Take some example of a morally wrong action, say the assassination of Martin Luther King. Can one really imagine that it's possible that another action that matched this one down to all of its qualitative natural properties—the facts about the victim, the intentions of the assassin, etc.—might be right? If you think this act was wrong, you seem bound to say that any act just like it (in relevant natural respects) is wrong as well.

So strong supervenience seems to hold between the ethical and the natural.¹⁰ That means that the natural properties instantiated in a situation metaphysically necessitate its ethical properties.¹¹ The challenge then is to explain why, or how the natural necessitates the ethical. If this were just a basic fact we could leave it as brute. But we have already argued above that modal facts aren't basic in this way, and so therefore we don't tolerate brute (metaphysical) necessities. Well, if it's not a brute fact, then what explains it?

Ethical naturalists have a ready answer, completely on a par with the answer to the corresponding question about mental supervenience given by the materialist: ethical properties are nothing over and above natural properties, either because they are identical to them or realized in them. As we've seen in the mind–body case, the claim that situations involving A properties are nothing over and above situations involving B properties is quite a good explanation for why instantiations of the latter necessitate instantiations of the former.

While ethical naturalists make the same explanatory move regarding supervenience on the non-normative that materialists make about supervenience on the non-mental, ethical non-naturalists diverge significantly from their counterparts in philosophy of mind, dualists (or anti-materialists). Anti-materialists agree with materialists that the only way to explain the supervenience of the mental on the physical would be to endorse the claim that the mental is nothing over and above the physical, and therefore is not fundamental. However, since they believe the mental is something over and above the physical—it is fundamental, or basic—they deny supervenience. This is the whole point of the appeal to zombies (along with the possibility of inverted qualia). Ethical non-naturalists, however, seem stuck with supervenience, for the reasons discussed above. Therefore, the challenge for them is how to explain it. If the ethical is something over and above the natural, why should ethical properties be necessitated by natural (non-ethical) properties?

⁹ For objections to supervenience, see Sturgeon (2009).

¹⁰ Interestingly, Blackburn (1993a, 1993b), who perhaps started the contemporary discussion of this issue, seemed to deny that strong supervenience held. He seemed to think that the issue that required explanation was why weak supervenience held, given that strong supervenience didn't. But as Dreier (1992) points out, this is mistaken. It is precisely that strong supervenience of the ethical on the natural that cries out for explanation.

¹¹ Again, as in the mind–body case, it's important that one include all of the relevant natural properties. Just as one can find apparent violations of supervenience with beliefs if one doesn't include the relevant surrounding physical facts, so too one might generate apparent counterexamples to ethical supervenience if one doesn't include all of the relevant surrounding natural facts.

So we are looking for a relevant difference between ethical properties and mental properties, one that explains why ethical non-naturalism can endorse supervenience without thereby committing itself to brute necessity, while the same is not true of anti-materialism about the mind. Well, to start, here is a difference. Everyone (except a few) accepts ethical supervenience while many do not accept mental supervenience. In fact, we can go further. In the ethical case, supervenience is held to be so obvious that its denial is often considered almost unintelligible. How could one call one action right and the other wrong if they do not differ in some natural way? What would that even mean? Whereas, not only do anti-materialists deny supervenience, it's often their denial of supervenience that drives their anti-materialism. That is, it isn't just that they deny supervenience because otherwise they would have to explain it or endorse brute necessity. Rather, based on conceivability considerations, they deny supervenience and then endorse anti-materialism as a consequence.

To see why this might matter, consider again the account of explanation given above. To explain some phenomenon is a matter of removing (relevant) alternatives; of showing why, given whatever is cited in the explanans, the explanandum *had* to obtain. In the mind-body case there really is a question whether or not supervenience holds; some would say this is *the* question in the debate. So if someone claims it does hold, they need to explain its holding, given that its not holding is a very live alternative. The options then are to either claim that the supervenience of the mental on the physical is a basic fact, or to explain it as a consequence of the mental being nothing over and above the physical. Since we've ruled out brute necessities, only the second explanation is available.

However, what we find regarding the supervenience of the ethical on the natural is that most philosophers can hardly find its denial intelligible. Indeed, there seems to be no real alternative to it; it just *has* to be the case. But if you already believe that some situation not only obtains, but it has to obtain, what is left to explain? Take whatever reasons you have for thinking it *has* to obtain and then that is your explanation for its obtaining.

While it would be nice if this were the end of the matter, it's not. As Jamie Dreier (unpublished manuscript) has emphasized, even if we grant that ethical supervenience holds by virtue of conceptual necessity, there is still a residual brute necessity involved.¹² Dreier formulates strong supervenience as follows:

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

This states that it is necessarily the case that for all properties (F) in the class of α properties, if an object instantiates it, then there is a property (G) in the class of β properties, such that the object instantiates it, and necessarily, anything instantiating G will instantiate F. This is in general what it is for the α properties to strongly supervene on

¹² For what follows I am indebted to Jamie Dreier and Tristram McPherson for pointing out inadequacies in my earlier response to the problem. I'm not sure they'll like what follows any better.

the β properties. He then claims that the first box, the one beginning the formula, is to be interpreted as conceptual necessity, and the second box as metaphysical necessity.¹³

The problem is that one still has to explain the inner box, the metaphysical necessity. The point is that one can't conceptually legislate that one property metaphysically necessitates another; there has to be an account of how that can be. For instance, to change an example of Dreier's, suppose I define "superbachelor" as an adult man who is necessarily unmarried. It's a perfectly consistent concept, but since it's unclear how a man *could be* necessarily unmarried, we have no reason to believe anyone could instantiate it. Similarly, while it may be a condition on a property's counting as an ethical property that it supervene on the relevant non-ethical properties, we still need to know how that can be explained.

As I said at the start of the chapter, I do think there is a challenge here for the ethical non-naturalist to meet, and I want to speculate about the best way to meet it. There are two reasons I am interested in this question: First, I'm very sympathetic to ethical non-naturalism. But more important, this metaphysical problem just doesn't seem intuitively to me like a good reason to abandon ethical non-naturalism, so there ought to be some way around it.

Let's return to our simplified framework for ethical properties described above. There are three ethical properties: right/ought-to-do, wrong/ought-not-to-do, and permissible/may-do-or-not-do. These properties apply to actions. Actions, let's say, are complexes of bodily movements and mental states, all of which count as non-normative for these purposes. So an action can be fully specified in non-normative terms. It then acquires (as it were, not intended temporally) an ethical value—right, wrong, or permissible—by virtue of the application to the fully specified action of an ethical evaluation. The evaluation procedure renders one of the three possible values as a function of the input parameters—the non-normative specification of the action—and the standard of evaluation that is determined by the objectively correct moral theory.

As a first stab at meeting the challenge, then, one might argue that ethical supervenience is logically/conceptually demanded by straightforward reflection on this structure. Supervenience of one set of properties on another can be put in terms of the former set being a function of the latter set. Well, so long as a standard of evaluation is univocal and objective, it's hard to see how the values it returns for action-types as inputs could be anything but a function of the inputs. There is, after all, nothing to influence the value returned for a given action input other than the nature of the input and the standard of evaluation.

Notice how the case of evaluative properties differs from properties the instantiation of which are determined by their causal/nomic relations to other properties. So, consider again the property dualist position on phenomenal consciousness. On this

¹³ In general I wouldn't claim that SS is a matter of conceptual necessity, but as it applies to the ethical case, where the α properties are the ethical properties and the β properties are the non-ethical ones, this is a way of capturing the sense of unintelligibility the denial of ethical supervenience seems to provoke in most people.

view phenomenal properties are basic properties that are tied to their physical/functional correlates by basic psycho-physical laws. According to such a law, when someone is in neurological state N_r , say, they will have a reddish phenomenal experience. As the basic psycho-physical laws could be different in other possible worlds (even ones in which the purely physical laws are the same), inverted qualia are possible, as are zombies (supposing the relevant basic psycho-physical laws don't obtain at all). However, unlike phenomenal properties on property dualist metaphysics, the instantiation of evaluative properties—whether basic or not—is determined by evaluative procedures, or standards, that are constitutive of those very properties; it is not a matter of lawful connection between the non-normative properties and the normative ones. And the fact that the evaluative procedure is constitutive of ethical properties helps not only to explain how mental and ethical properties differ, but also why ethical supervenience seems conceptually mandated; why its denial seems unintelligible.

To elaborate a bit, let me present a toy model. Suppose for simplicity's sake that straightforward Act Utilitarianism is the correct moral theory. So an action is right just in case it maximizes expected utility relative to alternatives, it's wrong just in case some alternative has higher expected utility, and permissible if no alternative has higher expected utility but some alternatives are equal to it. Is it logically, or conceptually possible that some action is both right and wrong? Clearly not. If rightness just is the first value that is determined according to the above standard, and wrongness just is the second value determined by the very same standard, then unless one action can both maximize and not maximize expected utility at the very same time—a logical impossibility—one action can't be both right and wrong at the same time. Notice that this impossibility—or better, its negation—doesn't count as a brute necessity, as it's underwritten by logic. But if the very same action can't, as a matter of logic, be both right and wrong, then how could two actions in distinct possible worlds that share all the features relevant to the ethical evaluation standard be right and wrong respectively? So have we just showed that ethical supervenience can be explained by logic after all?

I wish, but not quite. Earlier I emphasized that the way conceptual connections work to explain necessities has to do with the logical role of identity statements. I propose, then, that we accept the Act Utilitarian evaluation standard as underwriting identity statements like the following:

(R) Action A is right =_{def} A ought to be done by virtue of maximizing expected utility.¹⁴

Notice that it's not just it ought to be done *and* it maximizes utility. If that is all we meant, then we would be subject to the criticism that we're merely stipulating the necessary connection between being right and maximizing utility. Rather, the point

¹⁴ Or, A 's ethical value is "ought" as determined by the Act Utilitarian Ethical Evaluation Standard.

is that it's in the nature of rightness to be determined according to this standard—this is what it is to be right.¹⁵

Now (R) is a non-naturalist identity claim. A simpler, naturalist identity claim would be:

(R') Action A is right =_{def} A maximizes expected utility

Here is the problem. While (R'), together with the realization of expected utility in the relevant non-normative properties, explains how non-normative properties can necessitate normative ones, (R) at best explains how contradictory normative properties cannot be instantiated in non-normatively identical actions; it doesn't explain why any normative properties are instantiated at all.

There are two aspects to supervenience, whether it's mental or ethical: one that rules out "inversions" and one that rules out "zombies". In the mental case, an inversion would be a case in which there are two physically identical creatures and one is having a greenish experience and the other a reddish one, and a zombie case is one in which there are two physically identical creatures and one is conscious and the other isn't. In the ethical case, an inversion is a case of two non-normatively identical actions where one is right and the other wrong, and a moral zombie is a case of two actions that are non-normatively identical where one has a moral value and the other has no moral value at all. It seems to me that if we understand ethical properties as identical to the outputs of the correct, univocal, objective Ethical Evaluation Standard, then it's not hard to explain the absence of moral inversions.¹⁶ But ruling out moral zombies is another issue, and ethical supervenience does indeed rule them out, as attention to the supervenience formula described above makes clear.

For the non-naturalist—at least for this one—what is left over and above meeting the condition determined by the ethical evaluation standard (whether that be maximizing expected utility, satisfying the categorical imperative test, being chosen in the original position, etc.) is the normative force—that one *ought* or *ought not* do the action in question. As I see it, this is a fundamental property that is not realized in non-normative properties. By its nature its application is a function of the non-normative properties—what else is there to evaluate?—but that it applies at all in a situation cannot be derived from any representation of the situation's non-normative properties. Hence, by the framework developed above, if normative properties are

¹⁵ What I'm saying here bears a strong resemblance, it seems to me, to the argument Leary (forthcoming) makes. She wants to explain non-naturalist ethical supervenience by appealing to the "hybrid" essences of ethical properties, that they have both a natural and a non-natural character. I see this appeal to their being essentially the outputs of an evaluative procedure as a way of elaborating on this. However, as I make clear below, this doesn't completely solve the problem. I also see affinity between my account here and that of Enoch (2011).

¹⁶ Of course this assumes that there really is one, objective moral standard. I think this is safe to assume for our purposes because if one didn't believe this it's not clear why an ethical non-naturalist would hold ethical supervenience.

metaphysically necessitated by non-normative ones, this would constitute a brute necessity after all, of just the kind I wanted to rule out.

Earlier I said that bruteness was tolerable when dealing with a basic fact, and that basic facts were the kinds of facts that determined partitions in possible world space. The only bruteness attaching to all of possible world space I wanted to allow was the constraint of logic (along with conceptual structure if such there be). So, if this is right, then if normativity is basic, it can't be present in all possible worlds. I then believe we must give up on a part of ethical supervenience. Since it is not a matter of logic that an action instantiating such-and-such non-normative properties has an ethical value, it can't have it as a matter of metaphysical necessity. But once we see that all denying the metaphysical necessity amounts to is saying that there is something extra in the world, normative force, that is not guaranteed by the non-normative properties and logic, this shouldn't bother the ethical non-naturalist all that much.¹⁷ The position that normativity is something over and above the non-normative just brings that along with it. So long as moral inversions can be ruled out I think that's probably enough. Or so I think now.¹⁸

References

- Adams, R. M. 1987. *The Virtue of Faith and Other Essays in Philosophical Theology*. Oxford: Oxford University Press.
- Blackburn, S. 1993a. "Moral Realism." Reprinted in *Essays in Quasi-Realism*. Oxford: Oxford University Press, pp. 111–29.
- Blackburn, S. 1993b. "Supervenience Revisited." Reprinted in *Essays in Quasi-Realism*. Oxford: Oxford University Press, pp. 130–48.
- Chalmers, D. 1996. *The Conscious Mind*. Oxford: Oxford University Press.
- Chalmers, D. 1999. "Replies to Critics." *Philosophy and Phenomenological Research* 59: 473–93.
- Dreier, J. 1992. "The Supervenience Argument Against Moral Realism." *The Southern Journal of Philosophy* 30: 13–38.
- Dreier, J. (unpublished manuscript). "Is There a Supervenience Problem for Robust Moral Realism?"
- Enoch, D. 2011. *Taking Morality Seriously*. Oxford: Oxford University Press.
- Hume, D. 1998. *Dialogues Concerning Natural Religion*, ed. R. H. Popkin. Indianapolis: Hackett.
- Kim, J. 1993. *Supervenience and Mind*. Cambridge: Cambridge University Press.
- Leary, S. Forthcoming. "Non-Naturalism and Normative Necessities."
- Levine, J. 2001. *Purple Haze: The Puzzle of Consciousness*. Oxford: Oxford University Press.
- Levine, J. 2014. "Modality, Semantics, and Consciousness." *Philosophical Studies* 67: 775–84.

¹⁷ What I think should bother the ethical non-naturalist, however, is the epistemology. How can we have apparently a priori access to the fact that this is a world in which ethical values exist, and know their nature, if they are over and above the non-normative properties?

¹⁸ I would like to thank the following for very helpful discussion: Louise Antony, Jamie Dreier, Stephanie Leary, Ned Markosian, Tristram McPherson, Alejandro Perez Carballo, and the participants in the UMass Amherst Philosophy Brown Bag series, where an embryonic version of this paper was presented.

- Levine, J. 2016. "Naturalism and Dualism", in K. Clark (ed.), *The Blackwell Companion to Naturalism*. Malden, MA: Wiley-Blackwell, pp. 209–19.
- Lewis, D. 1986. *On the Plurality of Worlds*. Oxford: Blackwell.
- McPherson, T. 2015. "Supervenience in Ethics," in E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*. <<https://plato.stanford.edu/entries/supervenience-ethics/>>.
- Segal, G. 2000. *A Slim Book About Narrow Content*. Cambridge, MA: Bradford Books/MIT Press.
- Sturgeon, N. 2009. "Doubts about the Supervenience of the Evaluative," in R. Shafer-Landau (ed.), *Oxford Studies in Metaethics Volume Four*. Oxford: Oxford University Press, pp. 53–92.