Normative Commitments in Metanormative Theory[#][*]

Pekka Väyrynen

University of Leeds

1. Introduction

First-order normative theories concern what things are right and wrong, good and bad, just and unjust, rational and irrational, and so on, and why they are so. Metanormative theories concern the nature of first-order normative questions and claims. What is it to think that something is wrong? Are such judgments apt to be true or false? Is their truth a matter of objective fact? What would such facts be like? (And so on.) The two sorts of theories are widely treated as independent. For instance, suppose you advocate consequentialism in normative ethics. Then you think that any action is morally right just in case, and because, it maximises the good. It seems you can think this irrespective of whether your metanormative view of normative judgments is that they're some sort of conative states that cannot be true or false in any non-minimalist sense (expressivism), or that they're the sorts of things that can be robustly true or false and some of them are true (non-error-theoretic cognitivism), or even that they're truth-apt but systematically false (error theory). Similarly, you can endorse consequentialism irrespective of whether you think moral rightness is an objective property (whether natural or non-natural) or whether you think that knowledge of the truth of consequentialism is *a priori* or *a posteriori*. Accepting a particular first-order normative theory doesn't seem to force your metanormative hand.[1] Nor does your metanormative theory seem to force your normative hand. Most expressivists and error-theorists take their metanormative theories to be independent of whatever coherent answers to first-order normative questions we may

1   See e.g. Frankena (1951, pp. 45), Kagan (1998, pp. 5), Scanlon (1998, p. 59), and Hooker (2000, p. 15).

adopt.[2] Naturalist moral realism may pair most easily with consequentialism, but the pairing is by no means obligatory. Among non-naturalist moral realists, we find both both consequentialists and various kinds of non-consequentialists.[3]

First-order normative theory and metanormative theory are, then, widely seen as independent enterprises.[4] This paper argues that several debates in metanormative theory involve claims that have potentially surprising first-order normative implications. I first outline a general recipe for generating this result. I then apply this recipe to three debates in metaethics: the modal status of basic moral principles, normative vagueness and indeterminacy, and the determination of reference for normative terms. I argue that certain views on each issue carry first-order normative commitments in accordance with my recipe. In between, I address an important objection to two of these three applications of my general recipe.

2. A General Recipe

It's no news that some metanormative claims have normative implications. For instance, the kind of analytic naturalism according to which 'right' means 'maximises happiness' implies the truth of utilitarianism.[5] In these familiar cases the normative implications are direct and fairly obvious. Direct normative implications often lessen the credibility of those metanormative claims. For instance, even many utilitarians deny there are sound arguments for substantive normative conclusions like the truth of utilitarianism solely from premises about what normative terms mean (Moore 1903). I'll argue that some metanormative claims whose normative implications may not be immediately recognizable nonetheless have such implications, but needn't be less credible simply on that score.[6]

---

2   Such expressivists include Ayer (1936/1971, p. 105), Stevenson (1944, p. 110), Blackburn (1988, p. 367), and Gibbard (2008, p. 33). For error theory, see Mackie (1977, p. 16).

3   See e.g. Moore (1903), Ross (1930), and Scanlon (1998).

4   It may be difficult to say of some theories whether they are metanormative theories or normative theories (or both), and some families of theories may be open to be construed either way. Possible examples include Humean theories of reasons and constitutivist and constructivist theories of normative phenomena.

5   I'll use single quotes to denote terms or concepts, double quotes to introduce technical terms, and italics either for emphasis or to denote contents of sentences.

6   Many denials of independence are more sweeping than mine. Arguments from a particular metanormative theory to positive first-order conclusions in Hare (1981) and Korsgaard (2009) have

The metanormative claims on which I'll focus have normative implications because they make normative differences. The argument takes the following general form:

(M1) Factors picked out by some metanormative claims make a normative difference.

(M2) A factor F makes a normative difference only if F is normatively relevant.

(C1) So, factors picked out by some metanormative claims are normatively relevant.

(M3) The claim that F is normatively relevant is a normative claim.

(C2) So, some metanormative claims commit themselves to normative claims, to the effect that some factor is normatively relevant.

You can think of this as a general recipe for generating violations of the independence of metanormative theory from first-order normative theory. I don't take the recipe to generalise beyond the normative, for reasons that'll emerge in discussion. Examples of metanormative claims I take to satisfy (M1) will be contentious short of argument. (They won't carry normative implications in the direct way that analytic naturalism does.) So I must indulge your patience until I get to my applications of (M1). But to understand the recipe, we need to know what normative claims are and what it is to make a normative difference and be normatively relevant.

I know of no uncontentious analysis of what it is for a claim to be normative. But a leaf from Jonathan Dancy's account of normative facts gets us in the right ballpark. Dancy suggests that normative facts are explicit answers to the question what to do (or believe or feel or want). Some normative facts specify what to do fairly determinately, such as the fact that I ought to φ rather than ψ. Others answer that question in a less determinate way. The

---

widely been met with suspicion. Dworkin (1996) and Kramer (2009) argue that metanormative theories are also first-order theories, or else are to be settled on the basis of such. Berker (forthcoming) argues that to do first-order moral theory is also to do metaethics, but on the basis of controversial assumptions about the relation between grounding and normative ethics. Some writers seek to exclude some metanormative options on first-order grounds, but perhaps not always charitably; see e.g. Dancy (2004, ch. 4), Enoch (2011, ch. 2), and Parfit (2011, vol. 1, pp. 73-82 and vol. 2, pp. 277-88 and 457-9). More convincing, to my mind, are the views that any choice of a deontic logic rules out some logically tenable moral theories (Sayre-McCord 1986) and that statements of identity or reduction between normative properties and natural or supernatural properties are normative claims even if they're also metaphysical claims (Heathwood 2012).

fact that chocolate is good is clearly practically relevant, but doesn't specify just what difference it makes to what to do. In addition to simple normative facts like these, there are more complex normative facts that contain an atomic normative fact plus a reason why it holds. These can be expressed by claims like *That P makes it the case that you ought to act in way W, S ought to φ in virtue of the fact that P*, and so on. Such facts constitute explicit answers to both the question of what to do and the question why. It's crucial to distinguish normative facts from facts that are normatively relevant or significant – from facts that make a difference to what to do. That an action would please many and pain none is a non-normative fact. This fact, call it H, could be part of an answer to whether this action is the thing to do. But H isn't the same fact as the fact that H is relevant to what to do. The latter fact is a meta-fact about H. That meta-fact is normative. It's some such fact as that H is a reason to do the action, or that H makes it the case that one ought to do it, or that H is a reason why doing it would be good. (Dancy 2006: 136-8.)

I propose to understand normative claims on the same model, as claims that constitute answers (of varying determinacy in their content) to the question what to do (or believe, feel, or want). (At this level of precision, claims may be either propositions or (interpreted) sentences.) The claim that P is a reason to do an act is a normative claim because it's an answer to the question what to do. (It's a partial answer. Other partial answers are given by other reasons for and against the action.) P alone is only part of such an answer, and thus not itself a normative fact. When I defend (M3) below, I'll in effect propose that the claim that something is normatively relevant is a normative claim because it's a claim, about some other fact, that that fact makes a normative difference to what to do (or believe, feel, or want), and in that way constitutes a partial answer to the question what to do. This characterisation of normative claims runs in the material mode. It takes the contents of answers to the question what to do to be facts about what's right, what's good, what one ought or has reason to do, and so on. Ordinary normative thought and talk allows freedom of movement between claims about what's right and claims about what falls into the extension of 'right' (word or concept). But it's controversial whether these latter claims in the formal mode also count as normative claims. They have some claim to do so. That *x* falls into the extension of 'right' answers the question what to do. However, one might think that its license to answer that question derives from the feature of the truth predicate that *x is right* is true if and only if *x* is right, and *x*'s falling under 'right' isn't normatively relevant as such. These questions will arise in sections 4-6 below.

(M2) states a sufficient condition for normative relevance. That x is F is normatively relevant if x's being F makes a difference (in some non-causal way) to the normative features of x. So here I use "normative relevance" as an umbrella term for ways of making a normative difference, whatever that might include. The intuitive thought is just that if things differ only in normatively irrelevant ways, differential responses to them are unjustified, whereas if there are normative differences between things, then differential responses are justified. Consider a special case. To make an action wrong is to make a normative difference. If being F makes actions wrong, then (so far as that goes) something that's F wouldn't have been wrong had it not been F.[7] So being wrong making is a way of being normatively relevant; it's relevant to how normative features are distributed across things. But it's not the only way. That an action is F might be normatively relevant not by *making* it have a normative feature N but by *enabling* some other feature G to make it have N or by *intensifying* the contribution that F makes to N (Dancy 2004). Thus what makes a difference to whether x is wrong needn't be what explains why it is a fact that x is wrong. (That F enables G to make x wrong doesn't explain why x is wrong; what it helps to explain is the distinct normative fact that G makes x wrong.)

In what follows I'll apply (M1) to metanormative claims that make normative differences in none of these three ways, but in ways that aren't as immediately recognizable as normatively relevant. In fact, two of my three applications of my recipe will require that differences in the reference of normative terms and concepts are normative differences that fall under my recipe. This claim is controversial. But since the controversy doesn't affect my basic recipe as such but only its scope, I'll leave it until section 5. Here I'll instead say something a bit more general about how normative difference-making might be understood. It might be characterised in terms of counterfactuals, such as *If F hadn't obtained, then normative condition N wouldn't have obtained* and *If F hadn't obtained, N would still have occurred*. Such counterfactuals might not suffice for an adequate *analysis* of normative relevance.[8] But their (non-vacuous) truth might still be a (fallible) test for

---

7   Counterfactuals like these help to explain why we want to restrict 'F' to qualitative features. If Robbie the Rapacious Robber didn't exist, his wrongful robbing wouldn't exist. But what matters normatively to the wrongness of his actions is that they involve robbing, not that they involve Robbie. If rapacious robbing had occurred, but not by Robbie, it would still have been wrong.

8   Some tweaks will be forced by the distinction between derivative and non-derivative relevance. Suppose I've promised to give you an F (say, an apple). I give you x. Whether x is F makes a difference to whether I did as I ought: it makes a difference to whether I've fulfilled my promise. So the normative relevance of Fness derives from the fact that F figures in the content of my promise (where the fact that I promised to φ either is, or else derives its relevance from something that is, non-

whether F is normatively relevant to N (in the given circumstances). Or at least that's so under some such refinements as requiring the difference-making to be robust under some degree of abstraction or across a certain range of counterfactual variation. For instance, certain reasons to help someone who is a friend obtain would still obtain even if there wasn't a friendship but some other special relationship, but not otherwise. A different way to characterise normative difference-making is as follows.[9] Take a scenario S consisting of various facts. To determine whether a given fact F in the scenario makes a difference to a normative fact N in the scenario, compare any actual scenario S in which F is present with the scenario S* that results if we remove F from S. If there's no variation between S and S* with respect to N, then F is naturally described as not making a difference (with respect to N). If there's variation, F does make a difference. This is so whether the variation is that no normative condition at all holds in S* or that a normative condition N* holds which is distinct from N but has the same valence as N (say, that a state of affairs is less good or there's more reason to perform an action). So understood, difference-makers support the above counterfactuals and their status as at least a fallible test. Under either characterisation, factors that make a difference to what normative conditions hold are naturally described as normatively relevant, since they'll be relevant to how normative features are distributed across things. (Neither characterisation takes sides on whether normative claims come only in material or also formal modes.)

So much for (M2). Now consider (M3). What (M3) claims to be normative is the claim, about some fact, that that fact is normatively relevant. (Many normatively relevant facts are themselves non-normative.) Again, being wrong making is a way of being normatively relevant. (M3) is plausible in this special case. The claim that some fact is wrong making is a normative claim, whatever else (such as a metaphysical claim) it might be. Such a claim states that a certain fact stands in a certain normative relation to an act. I myself find it hard to believe that *Kicking puppies is wrong* is a normative claim but *Kicking puppies is wrong because of the pain it causes* isn't. What the second adds to the first is just a reason why kicking puppies is wrong – a kind of reason that can be used in a moral argument irrespective of one's views on the metaphysics of ethics. How is that not a normative claim? Similarly, moral principles can be thought of as identifying

---

derivatively normatively relevant).

9 This characterisation adapts the general template for accounts of causal difference-making proposed by Strevens (2008, p. 55) and the application of that template to define non-causally difference-making grounds in Krämer and Roski (2017). Its application will be sensitive to our background assumptions about what alterations in scenarios (including various counter-possible alterations) are coherent.

morally relevant factors (or, perhaps, factors that explain the moral features of things), and statements of moral principles are normative claims. *Acts are made wrong solely by failing to maximise general happiness* is exactly the sort of thing that a utilitarian moral theory says and which non-utilitarian theories are bound to deny. Nor does the special case of claims about what's right/wrong making look like an exception. Suppose that the fact that my promise wasn't given under coercion or fraud enables the fact that I promised to φ to be a normative reason for me to φ. That some factor enables another fact to be a reason to do something says, of that factor, that it's normatively relevant. Being told that some present fact is an enabler doesn't tell us anything very determinate about what to do, but it's nonetheless a partial answer to the question what to do. (M3) generalises this. The claim that something makes a normative difference (in whatever way such differences can be made) is a partial answer to the question what to do, and thus a normative claim.

What hangs on classifying the claim that something is normatively relevant as a normative rather than non-normative claim? Not much, one might say. It makes little difference in, say, biology whether the fact that certain (non-biological) facts make certain biological facts obtain is itself a biological fact. And even if claims about what makes biological facts obtain count as biological claims, one might think they're different from other biological claims. But examples like this don't convince me that it makes little difference whether we accept (M3). Even if the claim that something is normatively relevant is also a claim about (say) the metaphysics of normativity, it isn't only that.[10] Claims of normative relevance concerning what factors make things right and wrong or enable other factors to do so are needed to distinguish rival first-order normative theories. Theories may agree in their claims about what's right and wrong, and only differ in their claims about what makes them right and wrong. However things stand in biology, one would expect to distinguish normative theories by their normative content, not by some non-normative claims they make.

I also doubt that the biological (or the chemical, or the like) is the right analogy with the

---

10   Note that expressivists classify claims like *Kicking puppies is wrong because of the pain it causes* as normative claims. The moral sensibility of someone who thinks that kicking puppies is wrong because of the pain it causes is different from someone who thinks that kicking puppies is wrong because we don't like it (Blackburn 1988). But this latter point seems robust across cognitivist theories of normative judgment as well. I would thus be surprised if the status of such claims as normative were just an optional choice point for theories of normative judgment. Their status as normative strikes me as part of the subject matter.

normative in settling whether to accept (M3). The biological looks more like the aesthetic and the prudential. It might indeed make little difference whether the claim that some factor is prudentially or aesthetically relevant is itself a prudential or aesthetic claim. But the normative seems more like the causal in this respect. That something is causally relevant to an effect is a causal claim. That something is normatively relevant is, in the parallel way, a normative claim. By ascribing some or other kind of normative difference making, it constitutes an answer to the question what to do. For example, the claim that F enables R to be a reason to φ constitutes a partial such answer. It implies that there is a reason to φ when F holds. If the fact that a factor makes a certain kind of normative difference weren't itself a normative fact but only a normatively significant fact, it wouldn't constitute such an answer (however indeterminate). Whether we accept (M3) matters at least in this respect.

With my basic recipe on the table, I turn to three issues in metanormative theory: the modal status of basic normative principles, normative indeterminacy, and the determination of reference for normative terms. In each case I'll show that at least some theories about each issue conform to the recipe and thus have first-order normative implications.

3. The Modal Status of Normative Principles

What kind of necessity do normative principles have? Consider basic moral principles. Many think that, whichever these are, they hold by metaphysical necessity. But some think they hold as a matter of a weaker type of *normative* necessity (Fine 2002; Rosen forthcoming). Such principles are "fact-independent", in that they hold and would've held no matter what the world would've been like in wholly non-normative respects (Rosen forthcoming). If P1 is the basic moral principle (or the set of basic principles) holding in our world, then P1 would've been true no matter what we had thought, no matter how hard we had tried to falsify P1, no matter what the laws of nature had been, and so on. But no moral truths are metaphysically or otherwise absolutely necessary. In some worlds, some other fact-independent basic principle P2 holds. I'll argue that this view about the modal status of moral principles conforms to my recipe.

Views on the modal status of basic moral principles are metanormative. They concern a certain feature of basic principles and require no stand on which principles are basic. But the

8

metaphysical contingency of such principles would make first-order normative differences. Some things that are right in our (and many other) worlds aren't so in some other worlds even if those worlds are identical in all non-normative respects. If P2 doesn't hold in our world, acts that satisfy it won't always be right in our world. But if basic moral principles are metaphysically contingent, then acts satisfying P2 will always be right in some other worlds, no matter how the non-normative facts had been. So P2 is the basic principle of morality in some "counter-moral" worlds, just as some "counter-nomic" worlds that don't conform to the actual laws of nature exhibit some other laws. The metaphysical contingency of basic moral principles doesn't mean that just anything could (metaphysically) have been good or right (Bader 2017). It might be necessary that the basic normative principle is either P1 or P2 and nothing else, even if each of P1 and P2 holds where it does with mere metaphysical contingency. But it does mean that what's good or right might (metaphysically) not have been so.

The metaphysical contingency of basic moral principles is a witness to (M1) in my recipe. Let scenarios S and S* be identical in their qualitative non-normative properties, P1 the basic moral principle in S, and S* the result of removing P1 from S. Whether or not some other principle P2 must be substituted for P1 in S*, the subtraction of P1 means that normative properties are distributed differently in S and S*. So, given this test, the metaphysical contingency of P1 makes a normative difference. (That's so even if the metaphysical contingency of P1 isn't what *explains* the different normative facts in S and S*.) The rest of my recipe then entails that the claim that basic moral principles are metaphysically contingent commits us to certain first-order normative claims.[11] That's so even if the normative differences to which their metaphysical contingency commits us aren't immediately recognizable as normative claims.

To illustrate this more concretely, suppose deontology is true in our world but act-utilitarianism is true in some other worlds. (Constraints such as consistency and universalisability don't rule this out.) Act-utilitarianism might (metaphysically) have been true in our world. Knowing all this, I can ask: Why shouldn't we adopt act-utilitarianism instead? After all, we could accept and act on alternative moral principles which hold in counter-moral worlds. (The counter-nomic case is different: we cannot but obey the

---

11 One note of potential caution here is that which normative claims we regard as metaphysically contingent (or not) can make a difference to how to draw the distinction between the normative and the non-normative. For some examples of this (though rather different from mine), see Woods and Maguire (2017).

physical laws of our world.) And it's possible to do so on the basis of sincere and competent moral reasoning. Mistaken as they are, utilitarianism is what strikes my utilitarian friends as morally authoritative by their best lights. For better or worse, the metaphysical contingency view will reply that we should do what's right in the actual world, not what would be right if we were in a world where some alternative principle is true.

The point I want to make is that the metaphysical contingency of the basic moral principle isn't normatively irrelevant. If it were, then by (M2) this modal fact would make no normative difference. But it does make one. Different facts about what's right (with no difference in non-normative facts) hold depending on whether any such given metaphysical contingency obtains. That difference had also better be authoritative in the way normative truths are. If someone says that our best theory of what basic moral principles hold in our world is deontology, my utilitarian friends will offer reasons why utilitarianism is a morally better principle. They had better be wrong. For if utilitarianism were a morally better principle, how could deontology be the true morality in our world?[12] The general point holds irrespective of which principle holds in a given world with metaphysical contingency. (If there are reasons to think that deontology is a better morality, it seems false that utilitarianism is the best theory of the basic principle in the putative counter-moral world.) The basic principle had better be a principle which, given the alternatives, isn't subject to improvement and so merits our allegiance and compliance.

Given all this, (M3) is also plausible in this case. The view that basic moral principles are metaphysically contingent is committed to some first-order normative claims. One example would be the claim that the basic moral principle in our world is metaphysically contingent but not such that some alternative principle better merits our allegiance and compliance. Another example might be the claim that those non-normative facts which would be normatively significant under the alternative moral principles in fact are normatively irrelevant, and thus not what make right acts right.[13] To be sure, these normative claims don't constitute any very determinate answers to the question what to do. But not all normative claims do so anyway. Note that if the view that basic moral principles are metaphysically contingent does carry first-order normative implications, this doesn't by itself make the view less credible. Nothing in that claim shows that the basic principles that justifiably make a claim on our allegiance couldn't be different in different (non-

---

12 For a more extensive discussion of this kind of point, see Sayre-McCord (MS).

13 It's more controversial whether this is a normative claim.

normatively identical) worlds. Perhaps the normative differences made by the metaphysical contingency of those principles do have a sound, normatively non-arbitrary basis.[14] But first-order implications with a sound basis are still first-order implications.

4. Normative Indeterminacy

My two other applications of my general recipe concern metanormative claims that make a difference not to how normative properties are distributed but to the extensions of normative predicates. In this section I apply the recipe to the claim that the extensions of many normative predicates are *vague* or *indeterminate*. Is it morally permissible to abort a 24-week-old fetus? Is it morally permissible to save a friend at the expense of four strangers' lives? Cases like these might constitute borderline cases of moral permissibility. (Compare: Patrick Stewart clearly is bald, Ian McKellen clearly isn't, and in between there are borderline cases.) For illustrative purposes, we can think of vagueness as involving "semantic indecision". There are multiple precise extensions that are candidate referents for a word, but our linguistic conventions don't enforce a choice of any one of these sharpenings as the official referent (Lewis 1986, p. 213). So if 'morally permissible' is vague, there might well be properties P1 and P2 such that (i) they fit ordinary facts about the use of 'morally permissible' equally well but (ii) destroying a fetus at 24 weeks of gestation determinately has P1 and lacks P2. In that case, our linguistic conventions leave the extension of 'morally permissible' indeterminate between P1 and P2.

The claim that the extensions of normative predicates are indeterminate is a metanormative claim about the semantics of normative terms. One view in this debate is that in fact there's no indeterminacy despite semantic indecision. Further factors besides how a word is used in a linguistic community can play a role in determining its reference. For instance, "reference magnetism" is the idea that certain entities are intrinsically more eligible to be meant. They serve as "magnets" for expressions to refer to.[15] If some properties are "elite" in this way, a word can get a precise extension even if there are multiple precise extensions which fit facts about use equally well. Views vary on what – such as greater metaphysical fundamentality or objective naturalness – makes entities elite.

---

14 I assume throughout that normative truths, whether necessary or contingent, had better not be arbitrary.

15 For various versions of reference magnetism, see Hawthorne (2007), Sider (2011), Weatherson (2012), and Williams (2018). The idea of reference magnetism is often attributed to David Lewis (1983, 1984), but whether he endorsed it is controversial; see Weatherson (2012) and Schwartz (2014). The label "reference magnetism" was first used by Hodes (1984) for a related view.

Reference magnetism is a view in "metasemantics". It's invoked to explain in virtue of what words and sentences have the semantic properties they do, where what semantics aims at is a systematic characterisation of those semantic properties.[16]

Regarding 'morally permissible', suppose that a candidate referent P1 has whatever it takes to be more elite than P2, the only other contender.[17] According to reference magnetism, it follows that *Destroying a fetus at 24 weeks of gestation is morally permissible* is determinately true if and only if destroying a fetus at 24 weeks has P1. Whatever the merits of reference magnetism as a partial theory of how normative terms get their referents, this looks like a morally substantive implication. That P1 is a reference magnet for 'morally permissible' makes a difference to its extension. If P1 hadn't been elite, then different clams about what's morally permissible, and thus different answers to the question what to do, would have been true.

The conclusion that this metanormative claim carries first-order normative implications can be reached through a dilemma. Either being a reference magnet is a normative or a non-normative property. (On most views, being fundamental and being objectively more natural are non-normative properties.) If the claim that P1 is magnetic is normative, the claim *Our predicate 'morally permissible' refers to P1 because P1 is reference magnetic* commits us to a normative claim.[18] If that claim is non-normative, then it at least presupposes that being reference magnetic is normatively relevant. Differences in what falls and doesn't fall under 'morally permissible' make a difference to what to do. That P1 is magnetic doesn't itself explain why its magnetism makes that kind of normative difference. What's metaphysically elite needn't thereby be privileged in normative assessment (Eklund 2017, p. 30). But (M2) implies that if the fact that P1 is a reference magnet for 'morally permissible' does make a normative difference, it must be normatively relevant. This also seems independently plausible. If it's normatively arbitrary what properties are reference magnets, the truths that follow from such a reference assignment are normatively arbitrary as well, and thus not normatively relevant. But if the fact that P1 is a reference magnet for 'morally permissible' weren't normatively relevant, why should we be concerned with whether our actions have P1? What would give P1, and not P2, the claim to matter to the question what to do in this

---

16   See Schroeter and Schroeter (2017) for a useful survey of metasemantics in the context of metaethics.

17   Reference magnetism is applied to normative terms or concepts, in quite different ways, by van Roojen (2006), Dunaway and McPherson (2016), Suikkanen (2017), and Williams (2018).

18   The normative reference magnetism of Williams (2018) explicitly relies on certain substantive normative assumptions.

way?[19] So the metanormative claim *Our predicate 'morally permissible' refers to P1 because P1 is reference magnetic* has first-order implications.

According to reference magnetism, the extensions of vague normative predicates aren't affected by slight differences in usage. But my general style of argument may extend also to the view that the extensions of vague normative terms are "shifty".[20] On this view, 'morally permissible' lacks precise application conditions because its extension is highly sensitive to even slight differences in a linguistic community's usage.[21] The sentence of English *x is morally permissible* may refer to P1 when uttered by S but to P2 when uttered by S* who's otherwise just like S but belongs to a linguistic community that applies 'morally permissible' slightly differently from S. Thus different moral claims would be true if we used 'morally permissible' slightly differently. Such differences make a difference to the extensions of normative predicates. Looking at data about language use should then be a perfectly good method of moral deliberation insofar as it would be sufficient for us to learn which moral claims are true and thus constitute correct answers to the question what to do. Miriam Schoenfield points out that we're inclined to resist this thought (Schoenfield 2016, p. 266). My recipe promises to explain why. Non-normative facts about language use are fit to be used in deliberation about what to do only if they are normatively relevant. Unless they are normatively relevant, we have little assurance that when slight differences in use make a difference to the extension of 'morally permissible', this will reliably tracks facts about what's morally permissible. But if we might easily fail to learn substantive normative facts in learning facts about the extensions of normative terms, no wonder we're inclined to resist counting facts that might make a difference to their extensions as fit to be used in deliberation about what to do. The shifty view of normative language might address these worries by denying that factors that make a difference to the extensions of normative predicates are normatively relevant. But this reply requires rejecting plausible assumptions about the role their extensions can play in deliberation about what to do.

---

19 I focus on questions about our normative terms, and bracket the question why we should be concerned with our 'permissible' and not an alternative term 'permissible*' that plays the same (or similar enough) practical role as 'permissible' but picks out some different property P2. Eklund (2017) is an extensive discussion of complications raised by such alternative normative terms.

20 I lack the space to explore whether it extends also to some epistemic conceptions of normative vagueness.

21 Here I'll focus on differences that don't affect the practical role of 'permissible'. Differences in use which divest the term from its practical investment will generate differences in meaning. (Consider a community that uses 'permissible' to track their beer consumption.) But I don't take those to be relevant to my purposes.

My claims about these two views of normative vagueness run on a shared rationale. If some factor that plays a role in determining the reference of 'morally permissible' is normatively irrelevant, then facts about what falls into the extension of 'morally permissible' are to that extent normatively irrelevant as well. (They might be normatively irrelevant by, for instance, making normatively arbitrary differences to extensions.) But that would be a bad result. It would mean that truths about what's morally permissible wouldn't be fit to be used in answering the question what to do. It's much better to conclude that factors that make a difference to the extensions of normative terms only do so if they're normatively relevant, even if taking them to be relevant doesn't yield direct or highly determinate answers to the question what to do. This rationale will resurface in my argument in section 6 below. But first I'll address an important objection to applying my recipe to claims regarding the metasemantics of normative terms.

5. A Higher-Level Objection

My argument in section 4 requires that making a difference to the extension of a normative term or concept is making a normative difference. This is what my recipe requires to deliver the result that making a difference to the extension of a normative term or concept is a way of being normatively relevant. But there's a worry. Making a difference to the extension of a normative term by making a difference to the determination of its reference (in my main example, in the way that reference magnets do) isn't making a normative difference. That something makes this kind of difference to what falls under 'wrong' is a (metasemantic) claim about a normative term or concept that picks out wrongness, not a claim about wrongness. These are different subject matters. Outside the normative we recognize this. We sharply distinguish claims about tigers and claims about 'tiger', for instance. Why not do the same in the normative case?

A good way to bring about the objection is to consider what's needed to secure the license for the freedom of movement that ordinary practice tolerates between talk of what falls under 'right' and talk of what's right.[22] The acceptability of *p* and *'p' is true* go together non-accidentally, for the meaning of the truth predicate gives us the T-schema that *'p' is true* iff *p*. This equivalence of *p* and *'p' is true* is contingent, since *'p'* might have meant something other than *p*. What's needed to secure the license to move from the one to the other is just some or other mechanism that secures their actual equivalence. There's a

---

22   Thanks to Matti Eklund and Tristram McPherson for raising the kind of worries I try to address below.

general explanation that's compatible with all serious theories about how words and concepts get their referenc. Competent thinkers know that when they use 'F', 'F' predicates F-ness, and so you correctly predicate 'F' of *x* only if *x* is F. Nothing can be right but not in the extension of 'right' if we accept the T-schema.[23] Noting this fact seems sufficient to settle the question what to do in favor of φ-ing, if 'right' truly applies to φ-ing. (Add, if you like, that semantic facts about what falls under 'right' will be evidence of what's right.) We needn't suppose that the normative case is any different in order to explain why such semantic facts can be significant inputs to deliberation about what to do.[24]

The reply I'm inclined to favor is that when *p* is normative, our license to move between *'p' is true* and *p* does rely on certain normative presuppositions, and the normative is a special case because of the distinctive role of normative considerations in regulating deliberation and action. I won't be able to defend this claim fully here, but I'll motivate it with the following scenario. Suppose I learn that our best theory of what falls under 'right' in its moral uses in English ties the uses of 'right' in moral contexts to standards which I suspect to be flawed or defective in some relevant way.[25] Delicate questions arise regarding just how I might express this thought with my conceptual or linguistic resources, but the scenario seems intelligible.[26] (It's not automatic that English reliably tracks correct answers to the question what to do.) In such a case I would resist moving from *'x is right' is true* to *x is right,* in that I would resist accepting (in accordance with the role that 'right' plays in its normative uses) that *x* is the thing to do. I might resist on two different grounds. First, I might resist because I abandon the claim that *x is right* is true. I would do so on normative grounds, namely reasons to think that *x* isn't right. (In that case I'd have to think that what seemed our best semantic theory of what falls under 'right' in fact wasn't our best theory.) That doesn't seem sufficient to show that using the T-schema to move from the claim that it

---

23 Complications relating to semantic paradoxes and other hard cases can be bracketed as orthogonal here.

24 These assumptions about the deliberative role of normative terms may not hold across all views about the theoretical role of their extensions. For instance, different views about the relative normative importance of differently information-sensitive readings of 'ought' will recommend different views about the deliberative role of the extension of 'ought'.

25 The claim that the meaning of 'right' is a contingent fact about English allows this kind of scenario. But the scenario doesn't require 'right' to have a different meaning at the level of sense. Normative terms are plausibly semantically context-sensitive, and it's a highly salient possibility that in that case their extensions are underdetermined in the way the scenario requires. 'Right' can still well function as a term whose role is to settle questions about what to do, instead of being used to track some such descriptive facts as our beer consumption.

26 For a discussion of such delicacies (albeit one that brackets context-sensitivity), see Eklund (2017).

isn't the case that *x* is right to the claim that it isn't the case that *x is right* is true itself relies on normative presuppositions. But, second, I could instead grant that *x is right* is true – but only relative to poor standards, ones inappropriate to settling what to do. (This is compatible with thinking that given what 'right' *should* mean, or *should* refer to in this context, *x* isn't right.[27]) To get a feel for this possibility, suppose that 'right' refers to P1 because P1 is magnetic, but magnetism is determined at a level of metaphysical fundamentality which collapses some moral distinctions. In that case I could grant that *x is right* is true iff *x* has P1 but resist *accepting* that *x* is right, even though (by the T-schema) *x* is right. Accepting that *x* is right would commit me to doing *x*, and thus settle the question what to do in favor of *x*. But this isn't something to accept, since *x is right* comes out as true by ranking highly on standards (namely, P1) which I regard as inappropriate to settling what to do. The latter kind of reaction suggests that the standards that determine what falls under 'right' aren't supposed to be normatively irrelevant or arbitrary, given the distinctive role that normative considerations play in settling the question what to do. If those standards were irrelevant to settling what to do, then so would be the bearing of facts about what falls under 'right' to answers concerning what to do.

Given how I characterised normative claims in section 2, claims about what falls into the extension of a normative term or concept are normative claims if they constitute answers (however indeterminate) to the question what to do. My reply above doesn't strictly speaking require that claims about what falls into the extension of a normative term or concept themselves constitute answers to the question what to do. The reply only requires that for a metasemantic factor to make a difference to the extension of a normative term is a way of being normatively relevant and that the claim that something makes such a difference to what falls under a normative term is a normative claim. But my reply would be stronger if formal mode claims to the effect that something falls into the extension of a given normative term or concept were normative claims, whatever else they might be. So let's consider this briefly.

Normative claims in material and formal modes display similarities in their relations to normative reasons. We normally suppose there's a reason to do *x* either if *x* is right or if *x is right* is true. The latter fact doesn't itself seem like a reason to do *x*. But many think that the fact that *x* is right is also not a reason (or at least not a non-derivative reason) to do *x*. Or consider that we might count as normative those claims involving normative terms which

---

27 See Plunkett and Sundell (2013) for a discussion of these two kinds of "metalinguistic negotiation".

cannot be asserted consistently with normative error theory.[28] This test (no doubt fallible) correctly implies that *Stealing is wrong* and *Wrongness is instantiated* are normative claims. But it also counts *'Stealing is wrong' is true* as a normative claim. For error theorists would take that claim also to commit us to the existence of categorical reasons for action, and thus as much in error as *Stealing is wrong*. And we typically suppose we can settle both the question what's right and the question what falls under 'right' by considering reasons for and against particular things' being right. No doubt this supposition (just as the error theory test above) trades on the equivalence between *'p' is true* and *p*. But I suggested that in the normative case the freedom of movement between the two carries also normative presuppositions about their bearing on answering the question what to do and isn't exhausted by general facts about competence with the truth predicate. That said, some of these similarities may not hold across the board. For instance, even if a reason why x is right generally works also as a reason why *x is right* is true, the fact that x is right cannot make x right but might be thought to make it true that x is right. I've offered some reasons to think that formal mode claims about the extensions of normative terms and concepts count as normative claims, but clearly there's more to say about whether and when they do so.

6. Reference-Determination for Normative Terms

Earlier I applied my general recipe to certain metasemantic claims concerning the determination of reference for normative predicates which can be found in discussions of normative vagueness and indeterminacy. I'll now argue that my recipe applies also to the sort of causal theory of reference for normative terms that's usually associated with Cornell moral realism.[29] According to the causal theory of reference for a given term, the term has its reference determined by what its use is causally linked to in the appropriate way.[30] When such causal links are in place, our uses of the term are "causally regulated" by the property that's the referent. A causal metasemantics for the natural kind term 'water' explains why 'water' (as used by us) refers to H_2O by saying that H_2O is the substance which causally regulates our uses of 'water' (say, through communicative exchanges that generate appropriate causal links to initial dubbings of certain samples of liquid as 'water').[31]

According to causal theories of reference for normative predicates, normative predicates

---

28 Thanks to Alex Gregory (personal conversation) for this suggestion.

29 See e.g. Boyd (1988) and Brink (2001).

30 The appropriate way doesn't require speakers to associate any identifying description with the term.

31 See especially Putnam (1975), extending the causal theory of proper names in Kripke (1980).

also have their reference determined by what their use is appropriately causally linked to. Suppose there are two properties, P1 and P2, which fit the facts about our use of 'right' equally well. If P1 and P2 are properties such as maximising utility or meeting the test of the Categorical Imperative, we stand in causal links to each. But suppose that the property that in the relevant sense regulates our use of 'right' is P1. Then, according to the causal theory, 'right' picks out P1. In that case *x is right* is true iff *x* has P1. The explanation of why 'right' has this semantics would be in part that P1 is what causally regulates our use of 'right'. This metanormative factor makes a difference to the extension of 'right'. Had our use of 'right' been regulated by P2 instead, different things would have fallen under 'right'.[32] What I've suggested this to mean is that the causal theory identifies a metanormative factor that makes a normative difference.

The conclusion that this metanormative claim carries first-order normative implications can again be reached through a dilemma. Being a property that causally regulates our use of 'right' is either a normative or a non-normative property. If this property is normative, it's immediate that the claim *'Right' refers to P1 because our use of 'right' is causally regulated by P1* has normative implications. If it's a non-normative property, this metanormative claim implies that being such as to causally regulate our use of 'right' is normatively relevant. Causally regulating our use of 'right' makes a difference to what falls under 'right', and insofar as this is a normative difference, the fact that P1 regulates our use of 'right' is normatively relevant. The difference is makes looks normative. Note that it seems perfectly legitimate to ask why, given that our use of 'right' is causally regulated by P1, we ought to be concerned with whether our actions have P1 rather than P2.[33] If the fact that P1 plays this causal role made no difference to how to answer the question what to do, why should we be concerned about this? Since we normally do care about whether our actions are right, that an action has P1 is relevant to answering the question what to do. But in that case such a fact is normatively relevant, and I've suggested those are normative claims. So, either way, the metanormative claim *'Right' refers to P1 because our use of 'right' is causally regulated by P1* carries first-order normative implications.[34]

---

32 Again we must take care with the sense in which this means that 'right' would mean something different.

33 Similar points are made by Eklund (2012) and Plunkett and Sundell (2013).

34 There's an interesting question whether my recipe applies also to conceptual role theories, on which the reference of normative predicates such as 'right' is determined by what property best fits the conceptual role of 'right', which we may take to consist in its role in practical reasoning (Wedgwood 2001; Chrisman 2015). I lack the space to discuss the complications raised by this case. But see Eklund (2017, pp. 54-9).

Some might conclude from the above argument that causal theories of reference for normative terms should be rejected. Here's an example to illustrate the objection:

(O1) Helping the sick falls into the extension of 'right'.

(O2) Our use of 'right' is causally regulated by property P1.

(O3) According to causal theories of reference, (O1) holds in virtue of (O2).

(O4) (O2) as such is normatively irrelevant.

(O5) If X is normatively irrelevant, and Y obtains in virtue of X, then Y is normatively irrelevant.[35]

(C) So, if causal theories of reference are correct, (O1) is normatively irrelevant.

Nothing in causal theories of reference as such yields reasons to help the sick. A causally more apt regulator of our uses of normative terms needn't be correspondingly apt for normative assessment. This seems to make claims about what falls into the extensions of our normative terms inapt to settling what to do, and thus normatively irrelevant. But if helping the sick does fall into the extension of 'right', that's not normatively irrelevant. (Or so I've suggested.) Causal theories of reference for normative terms are therefore inadequate to their subject matter.

This objection fails. Causal theories can take the reference of a normative term to be constrained epistemically, by whether applications of the term enable our beliefs increasingly to approximate the relevant normative truths (at least provided our initial background theories are relevantly approximately true). Boyd, for instance, holds that 'good' refers to moral goodness only if moral goodness regulates the use of 'good' in moral and non-moral reasoning so as to provide a basis for the growth of moral knowledge (Boyd 1988, p. 201).[36] If our use of 'good' in reasoning doesn't help us approximate truths about moral goodness, our use is causally regulated not by moral goodness but some other property. So 'good' couldn't easily have referred to something other than what it actually refers to.[37] This also means that claims about what property causally regulates our use of 'good' make normative assumptions. For instance, if P1 is that property, this has

---

35 This is the "Principle of Irrelevancy Transmission" from Killoren (2016, p. 171).

36 As Boyd puts it, a term *t* refers to a kind (property, relation) *k* (in nondegenerate cases) just in case "the socially coordinated use of *t* provides significant epistemic access to *k*, and not to other kinds" (Boyd 1988, p. 195).

37 This point bears also on the Moral Twin Earth objection to naturalist realism. See Väyrynen (2018).

implications for what moral goodness consists in. If our assumptions about what moral goodness consists in are false, this will infect our assumptions about what property regulates our use of 'good'. The causal theory of reference for normative terms can thus agree that it carries first-order normative implications but claim that it needn't be less credible just because it does so. This matters. Many discussions seem to miss the point that causal theories can agree that the question of which property relevantly causally regulates our use of a normative term cannot be answered without raising the normative question of what the normative facts are.

7. Conclusion

I've argued that certain metanormative claims carry commitments to first-order normative claims. The novelty of my argument doesn't lie in this claim as such, but in the general recipe I offer for generating instances of such dependence and in the specific instances of the recipe which I defend. The cases I offer comprise certain views in debates on the modal status of normative principles and the metasemantics of normative terms. These views make differences which, while they may not be immediately recognizable as normative, I've argued to count as normative implications all the same. Carrying those implications needn't, however, make the metanormative views that carry them thereby less credible. I'll have to leave it an open question where else in metanormative theory my general recipe applies.

If my arguments are on the right track, this has ramifications for doing normative and metanormative theory. On the one hand, particular normative theories will rule out certain metanormative views, namely those whose normative implications are incompatible with the first-order theories in question. On the other hand, certain metanormative views will rule out certain first-order normative views, by implying certain others. Insofar as each sort of theories deploy the method of wide reflective equilibrium, in either case we'll need to consider whether to adjust the normative theories or the metanormative theories in question, and on what grounds. Either way, neither normative nor metanormative theorists can decide in advance to restrict their attention solely to their proprietary questions. Sometimes this will be safe, but sometimes it won't. The relevant safety checks can hopefully be informed by the general recipe I've offered for determining whether a metanormative view has first-order normative implications.

References

Ayer, A. J. (1936/1971) *Language, Truth and Logic* (Harmondsworth: Penguin).

Bader, Ralf (2017) 'The Grounding Argument Against Non-Reductive Moral Realism', *Oxford Studies in Metaethics* 12, pp. 106–34.

Berker, Selim (forthcoming) 'The Unity of Grounding', *Mind*.

Blackburn, Simon (1988) 'How to Be an Ethical Anti-Realist', *Midwest Studies in Philosophy* 12, pp. 361–75.

Boyd, Richard (1988) 'How to Be a Moral Realist', in *Essays on Moral Realism*, Geoff Sayre-McCord (ed.) (Ithaca, NY: Cornell University Press), pp. 181–228.

Brink, David (2001) 'Realism, Naturalism, and Moral Semantics', *Social Philosophy and Policy* 18, pp. 154–75.

Chrisman, Matthew (2015) *The Meaning of Ought: Beyond Descriptivism and Expressivism* (Oxford: Oxford University Press).

Dancy, Jonathan (2004) *Ethics without Principles* (Oxford: Clarendon Press).

Dancy, Jonathan (2006) 'Nonnaturalism', in *The Oxford Handbook of Ethical Theory*, David Copp (ed.) (Oxford: Oxford University Press), pp. 122–45.

Dunaway, Billy and McPherson, Tristram (2016) 'Reference Magnetism as a Solution to the Moral Twin Earth Problem', *Ergo* 3, pp. 639–79.

Dworkin, Ronald (1996) 'Objectivity and Truth: You'd Better Believe It', *Philosophy and Public Affairs* 25, pp. 87–139.

Eklund, Matti (2012) 'Alternative Normative Concepts', *Analytic Philosophy* 53, pp. 139–57.

Eklund, Matti (2017) *Choosing Normative Concepts* (Oxford: Oxford University Press).

Enoch, David (2011) *Taking Morality Seriously* (Oxford: Oxford University Press).

Fine, Kit (2002) 'Varieties of Necessity', in *Conceivability and Possibility*, Tamar Szabó Gendler and John Hawthorne (eds.) (Oxford: Oxford University Press), pp. 253–81.

Frankena, William K. (1951) 'Main Trends in Recent Philosophy: Moral Philosophy at Mid-Century', *Philosophical Review* 60, pp. 44–55.

Gibbard, Allan (2008) *Reconciling Our Aims: In Search of Bases for Ethics* (Oxford: Oxford University Press).

Hare, R. M. (1981) *Moral Thinking* (Oxford: Clarendon Press).

Hawthorne, John (2007) 'Craziness and Metasemantics', *Philosophical Review* 116, pp. 427–40.

Heathwood, Chris (2012) 'Could Morality Have a Source?', *Journal of Ethics and Social Philosophy* 6 (2), pp. 1–19.

Hodes, Harold (1984) 'Logicism and the Ontological Commitments of Arithmetic', *Journal of Philosophy* 81, pp. 123–49.

Hooker, Brad (2010) *Ideal Code, Real World* (Oxford: Clarendon Press).

Kagan, Shelly (1998) *Normative Ethics* (Boulder, CO: Westview Press).

Killoren, David (2016) 'Why Care about Moral Fixed Points?', *Analytic Philosophy* 57, pp. 165–73.

Korsgaard, Christine (2009) *Self-Constitution: Agency, Identity, and Integrity* (Oxford: Oxford University Press).

Kramer, Matthew (2009) *Moral Realism as a Moral Doctrine* (Oxford: Blackwell).

Krämer, Stephan and Roski, Stefan (2017) 'Difference-Making Grounds', *Philosophical Studies* 174, pp. 1191–1215.

Kripke, Saul (1980) *Naming and Necessity* (Oxford: Blackwell).

Lewis, David (1983) 'New Work for a Theory of Universals', *Australasian Journal of Philosophy* 61, pp. 343–77.

Lewis, David (1984) 'Putnam's Paradox', *Australasian Journal of Philosophy* 62, pp. 221–36.

Lewis, David (1986) *On the Plurality of Worlds* (Oxford: Blackwell).

Mackie, J. L. (1977) *Inventing Right and Wrong* (Harmondsworth: Penguin).

Parfit, Derek (2011) *On What Matters*, volumes 1–2 (Oxford: Clarendon Press).

Plunkett, David and Sundell, Tim (2013) 'Disagreement and the Semantics of Normative and Evaluative Terms', *Philosophers' Imprint* 13 (23), pp. 1–37.

Putnam, Hilary (1975) 'The Meaning of "Meaning"', *Midwest Studies in the Philosophy of Science* 7, pp. 131–93.

Rosen, Gideon (forthcoming) 'Normative Necessity', in *Metaphysics, Meaning, and Modality: Themes from Kit Fine*, Mircea Dumitru (ed.) (Oxford: Oxford University Press).

Ross, W. D. (1930) *The Right and the Good* (Oxford: Clarendon Press).

Sayre-McCord, Geoff (1986) 'Deontic Logic and the Priority of Moral Theory', *Noûs* 20, pp. 179–97.

Sayre-McCord, Geoff (MS) 'On a Theory of a Better Moral Theory and a Better Theory of Morality', unpublished.

Scanlon, T. M. (1998) *What We Owe to Each Other* (Cambridge, MA: Harvard University Press).

Schroeter, Laura and Schroeter, François (2017) 'Metasemantics and Metaethics', in *The Routledge Handbook of Metaethics*, Tristram McPherson and David Plunkett (eds.)

(London: Routledge), pp. 519–35.

Schwarz, Wolfgang (2014) 'Against Magnetism', *Australasian Journal of Philosophy* 92, pp. 17–36.

Schoenfield, Miriam (2016) 'Moral Vagueness Is Ontic Vagueness', *Ethics* 126, pp. 257–82.

Sider, Theodore (2011) *Writing the Book of the World* (Oxford: Oxford University Press).

Stevenson, C. L. (1944) *Ethics and Language* (New Haven: Yale University Press).

Strevens, Michael (2008) *Depth: An Account of Scientific Explanation* (Cambridge, MA: Harvard University Press).

Suikkanen, Jussi (2017) 'Non-Naturalism and Reference', *Journal of Ethics and Social Philosophy* 11 (2), pp. 1–25.

van Roojen, Mark (2006) 'Knowing Enough to Disagree: A New Reply to the Moral Twin-Earth Argument', *Oxford Studies in Metaethics* 1, pp. 161–194.

Väyrynen, Pekka (2018) 'A Simple Escape from Moral Twin Earth', *Thought* 7, pp. 109–18.

Weatherson, Brian (2012) 'The Role of Naturalness in Lewis's Theory of Meaning', *Journal for the History of Analytic Philosophy* 1 (10), pp. 1–19.

Wedgwood, Ralph (2001) 'Conceptual Role Semantics for Moral Terms', *Philosophical Review* 110, pp. 1–30.

Williams, J. Robert G. (2018) 'Normative Reference Magnets', *Philosophical Review* 127, pp. 41–71.

Woods, Jack and Maguire, Barry (2017) 'Model Theory, Hume's Dictum, and the Priority of the Ethical', *Ergo* 4, pp. 419–40.