

Normative Concepts

Matti Eklund

matti.eklund@filosofi.uu.se

for Oxford Handbook in Ethical Theory, edited by David Copp and Connie Rosati

1. Introduction

My topic is normative concepts. Before proceeding to more substantive questions, let me say something what is meant by “normative” here and what is meant by “concept”.

Let me start with the “normative”. The label “normative” has become quite widely used, and I myself find it useful. But it has come to have quite a few different uses. Some clarificatory remarks are in order.

First, there is a usage of “normative” on which the label applies to all broadly evaluative, deontic, prescriptive concepts, and there is a narrower usage on which a deontic concept like RIGHT is normative but evaluative concepts like GOOD are not. I will here use “normative” broadly. In part this is just a terminological decision. But there are nearby substantive issues. Is the class of the “normative” in the broad sense unified enough that we can reasonably expect unified accounts of it? Or is it more fruitful to theorize in terms of, for example, the deontic and the evaluative, respectively?

Second, “normative” is wider than merely what pertains to ethics and morality. Moral concepts would be regarded as normative, but there are arguably normative concepts that pertain to some other domain (like aesthetics) or that simply are more general (witness the last “ought” in “I know I *morally* ought to do it, but *ought* I really to do it?”).

Third, there is a distinction between *authoritative* and *merely formal* normativity, and sometimes “normative” is used for both while sometimes it is used only for the so-called authoritative kind. Formal normativity is the normativity at issue when something is said about how a player “should” move in connection with chess, speaking of what is instrumental for winning, and when something is said about what utensil one “ought” to use, speaking of what accords with rules of etiquette. Authoritative normativity is, intuitively, paradigmatically on display in connection with morality or in connection with questions of what all things considered ought to be done. As I will return to, it is difficult to provide a helpful characterization of the distinction, but it is generally agreed that there is a distinction of some

kind here. Sometimes the label “normative” is reserved only for what I have called the authoritatively normative; but I will keep using the label broadly.

Fourth, it is common, and arguably useful, to distinguish between *thin* and *thick* normative concepts. Among paradigm examples of thin concepts are RIGHT, GOOD, OUGHT, and REASON (meaning: the concepts the words “right”, etc., express in contexts of special importance to metanormative inquiry). Among paradigm examples of thick concepts are BRUTAL, GENEROUS, RUDE, and COURAGEOUS. The general idea is that thin concepts in some way have less descriptive, or less specific content than thick concepts do. Again it is difficult to give a satisfactory account of the supposed distinction, but it is widely agreed that there is a distinction here.

In connection with the use of the label “normative” I should also stress that focus on the broad category of the normative is a relatively recent phenomenon. Proponents of the view that this category is a proper one to focus on also often hold that instead of focusing specifically on ethics and morality, as within much traditional metaethics, we should attend more generally to *metanormative inquiry*. To stress, this relies on the theoretical assumption that the normative is a useful category for theorizing. I am attracted to this view myself but I will not defend it here –and it is by no means uncontroversial.

Let me now turn now to the notion of a *concept*. One must distinguish between *properties* on the one hand and *representations* such as concepts and linguistic expressions on the other. Properties are had, or instantiated, by entities. Concepts and linguistic expressions are rather true of – *correctly represent, apply to* – entities. If Socrates *has* the property of being wise, then the concept WISE *applies* to him. One might of course for theoretical reasons choose to identify concepts and properties, but *prima facie* there is an obvious distinction here, and for present purposes I will simply assume that appearances do not deceive.

Concepts, like linguistic expressions, are representational. Roughly speaking, concepts, as I will talk of them here, are the meanings of linguistic expressions. But it should not be assumed that there is a simple relation between expressions and concepts. A given expression may express different concepts on different occasions of use. And in principle we may have and use concepts that are never what is semantically expressed by any of our expressions on any occasion of use. That said, I will out of convenience often go back and forth between speaking of meanings and speaking of concepts.

Sometimes one speaks of concepts a different way, as mental counterparts of words. Thus understood, concepts are not themselves meanings any more than words are. Instead, concepts *have* meanings. One problem regarding talk of concepts in philosophical discussions

is that such talk can be ambiguous in the way just described. I will have reason to return to this.

Given the way the label “concept” is standardly used in philosophy, concepts must also be distinguished from what are sometimes called *conceptions*. Suppose you and I associate dogs with very different things. You, having been around friendly, cuddly dogs associate dogs with something positive; I, having had traumatizing experiences, think of dogs as scary. In some sense we associate different theories with dogs. We have different *conceptions* of dogs. For all this it can be that we have and employ the same *concept* DOG, in the relevant sense. For example, my thought *dogs are dangerous* and your thought *dogs are dangerous* may still have the same truth conditions. Compare: just as you and I can use the word “dog” with the same meaning despite associating different conceptions with it, we can in principle use the same concept DOG despite our different conceptions. This is not to say that we do actually have and employ the same concept DOG. The point is just that the mere fact that we have different conceptions does not entail that we employ different concepts.

Given the distinction between properties and concepts, one must distinguish between talk of normative properties on the one hand and talk of normative concepts on the other. The question arises what “normative” means as applied to properties and as applied to representations, respectively. To say that a property is normative sounds like it is related to its being something that ought to be promoted (or dispromoted, in the case of negatively valenced normative properties). But to say that a word or concept is normative is not supposed to convey anything about its being worthy of being promoted. This is of course not unique to the normative, but affects talk of concepts and talk of properties generally. To say that a property is, for example, a physical property is different from saying that a concept is a physical concept.

Distinguishing clearly between the issue of normative representations and normativity in the world also helps make plain that it is possible that there really is only normativity of one of these kinds. It could be that while some representations are reasonably called normative, the things in the world which they represent are not. There is a concept RIGHT, which is normative, and this concept stands for a property, rightness, but this property is not itself normative. Another possibility is that the only things reasonably deemed normative are worldly items. Maybe the *best* sense that can be made of “normative” as applied to representations is in terms of whether the representation simply tends to be used to recommend (or disrecommend, in the case of the negative), but then the representation itself would not be normative. (Compare: it may be that because people generally like kittens,

saying that something is a kitten serves to recommend it, but that does not make the expression “kitten” itself normative.)

Another possibility is that there is both representational and worldly normativity, but some normative concepts stand for properties that are not normative, and some normative properties are properties we have non-normative concepts for. I will get back to such possibilities.

2. Moore’s open question argument

Perhaps the most famous argument pertaining to normative concepts is G.E. Moore’s *open question argument* (1903). In a central part of the argument, Moore argues, plausibly, that when one compares a sentence using “good” and a sentence using a purported synonym using only non-normative terms, these sentences are not synonymous. The concept expressed by “good” and the concept expressed by the purported synonym do not have the same content. For example, consider the view that “good” means the same as “desired to be desired”. Then note that while the question “If something desired to be desired, is it desired to be desired?” is trivial, the question “If something desired to be desired, is it good?” is not.

The conclusion Moore ultimately wants to draw is that the *property* of being good is not definable. The property that “good” (and the concept good) stands for is not identical to the property that the purported synonym (and the concept it expresses) stands for, and the way this generalizes is supposed to entail the indefinability of the property.¹ A common, and well-taken, criticism of Moore is that even if Moore is right in what he says about expressions and concepts, his conclusions about properties just do not seem to follow. “Water” and “H₂O” do not mean the same, yet they can stand for the same property. Why can not the same be true of “good” and a descriptive counterpart?² Despite the problems with Moore’s argument as it stands, some later theorists have held that improved version of the argument may work.³ And Moore’s claim about the property of being good can certainly be true and plausible even if his argument for it should not persuade.

Thus far I have gone through Moore’s open question argument and the standard reaction to this argument. What, if anything, does any of this show about normative expressions and concepts more generally? A natural generalization of Moore’s claim about

¹ Of course, the argument as presented leaves open that the property of being good is definable in terms of other normative properties. But if it can, then Moore can instead simply focus on these other normative properties, so this is not a major problem for Moore’s argument.

² See e.g. Sturgeon (2003).

³ See e.g. Dancy (2006) and FitzPatrick (2011).

“good” may be that no normative expression is synonymous with a non-normative expression. This general claim is apt to sound trivial. Synonymy is sameness of meaning, and if an expression’s normativity is a matter of its meaning, then trivially normative and non-normative expressions are never synonymous. But there are two responses to this in turn. First, it is not trivial that the normativity of an expression has to do with its meaning. It could attach to the expression or concept some other way. Maybe, for example, the normativity of an expression is a matter of its *lexical effects*, where these are effects of the use of the expression which do not have to do directly with its semantics or its pragmatics, and insofar as one can speak of the normativity of concepts that is a matter of the counterpart of lexical effects in thoughts.⁴ Second, even while taking on board that non-normative expressions trivially never can have the same meaning or content as normative expressions, one can treat it as an open possibility that paradigmatic classifications are mistaken – a paradigmatic example of a purportedly normative expression can turn out to be synonymous with a paradigmatic purportedly non-normative expression.

The remarks in the above paragraph concern expressions. How do they transpose to the case of concepts? Insofar as one thinks of concepts as mental counterparts of words, they transpose straightforwardly. Just as words have features that are not features of their meanings, so do concepts. But insofar as one instead thinks of concepts as themselves being meanings or contents, any normativity pertaining to a concept itself must have to do with meaning or content.

One natural generalization of Moore’s claim that the property of being good is not identical to any property a non-normative expression stands for is to say that no property that a normative expression stands for can be identical to any property a non-normative expression stands for. I will get back to issues like this shortly. For now, let me just note that the problem with Moore’s argument certainly also affects a corresponding argument for this more general claim.

A different broadly Moorean claim in the same ballpark as the claim about the property of being good is that no normative property is identical to any non-normative property. This claim is apt to sound trivial: if $F=G$ then F and G share all their properties, so of course if one is normative the other is too. One reaction is again that even if this indeed is trivial, it is an open possibility that our views on what properties are normative and non-normative may be revised. If it seems to turn out that seemingly normative property F is

⁴ See Cappelen (2018), ch. 11, for more on lexical effects, albeit not in connection with the specific topic of normativity.

identical with seemingly non-normative property G, then either F isn't normative or G isn't non-normative.

3. What makes a concept normative?

One natural question to ask regarding normative concepts is: what makes a concept normative? Note that to answer this question it is not sufficient to give accounts of how specific normative concepts work. Even given correct, informative accounts of GOOD, BAD, RIGHT, OUGHT, etc., the general question of what (if anything) makes these concepts normative remains.

The notion of being a normative concept is something of a term of art, and not only because “normative” has a special use in philosophy: even given general clarity regarding what “normative” means one may wonder about the locution “normative concept”. Given that it is, it may seem silly to argue over what it really is for a concept to be normative: can we not stipulate a meaning for “normative concept” and get it over with? But I will here proceed on the assumption that there is a reasonable question here.

A first suggestion regarding what makes a concept normative is that a concept is normative by virtue of standing for a normative property. This would among other things entail that all concepts that stand for normative properties are normative. But that consequence seems false. Consider the following possibility.⁵ The Martians come to Earth. They recognize that humans in some parts of the world use “right” and take an interest in this. They introduce the word “thgir” into their language by the stipulation that it stand for whatever property “right” stands for. They then come to employ the concept THGIR. According to the suggestion at issue, it is sufficient for THGIR to be normative that it stands for the normative property of being right. But it seems straightforward to elaborate on the case so that THGIR does not intuitively seem normative. For example, the Martians may not in any way use THGIR to recommend or disrecommend or otherwise evaluate. They might take a purely clinical, descriptive interest in what is “thgir” and not. (If it helps jog intuitions, one might even add that the Martians seem incapable of normative thinking as we know it, and seem to have no normative concepts of their own.)

Notice that the “thgir” example also seems to refute what I described as the natural generalization of Moore’s claim regarding the property of being good. The generalization was that no property that normative expression stands for can be identical to any property a non-

⁵ The example that follows is from Eklund (2017), pp. 75ff. Other theorists, for example Wedgwood (2018), have introduced and made use of similar examples.

normative expression stands for. “Right” is normative and “thgir” is non-normative, but they stand for the same property. Hence the generalization is false. And the point is general, and could of course be made also with, e.g., a non-normative expression “doog” standing for the same property that “good” stands for. Faced with this point, one might try to formulate a revised claim that avoids these counterexamples. One might suggest that “thgir” gets its reference *via* the reference of “right” and may be said to refer to the property it refers to only derivatively. The argument I have presented does not by itself rule out the possibility that no property a normative expression stands for can be identical to any that property that a non-normative expression *non-derivatively* stands for.⁶ Anyone inclined to appeal to this must make clear what the relevant notion of being derivative amounts to.

Appeal to examples like THGIR when arguing about what makes a concept normative is apt to provoke the kind of reaction mentioned above. Can we not just *decide* to use the term of art “normative” in such a way that a concept is normative just in case it stands for a normative property? But there are two things to say in response. First, while the notion of a normative concept may be technical, the general practice of classifying concepts as being of different kinds is not. And in general, we do not say that a concept is of given kind simply in virtue of ascribing a property of a corresponding kind. Compare: even on the assumption that the property of being in pain is identical to the property of having one’s C-fibers firing, so a psychological or mental property is identical to a physical property, the concept PAIN is hardly a physical concept and the concept C-FIBERS FIRING is hardly a psychological concept. Second, even while allowing that we are free to stipulate a meaning for “normative concept”, it remains that simply choosing to align “normative concept” and “concept that ascribes a normative property” seems to amount to not taking the concept/property distinction very seriously.

The example of THGIR then arguably shows that a concept can stand for a normative property and yet fail to itself be normative. Can a concept be normative even while standing for a non-normative property? Arguably yes. Consider thick concepts. Some thick concepts seem somehow to rely on in some ways flawed moral assumptions and thus to be normatively *objectionable*.⁷ For example, many concepts pertaining to sexual morality, such as for example LEWD, (SEXUALLY) PERVERTED, CHASTE, and PROMISCUOUS seem to rely on mistaken assumptions of a conservative or reactionary nature. Yet these concepts do stand for properties, and those properties are instantiated. Some people really are lewd, chaste,

⁶ Roberts (2013) suggests taking this kind of route.

⁷ For discussions of objectionable thick concepts, see Väyrynen (2013) and Eklund (2011, 2017).

promiscuous, etc. Are the properties these concepts stand for normative or not? I would say: they are not – and a way to describe the issue with these problematic concepts is that they present non-normative properties in a certain normative light.

A second suggestion regarding what makes a concept normative is to appeal to relations between concepts. Maybe a concept is normative by virtue of standing in the right sort of conceptual relation to some paradigmatic normative concept like REASON or OUGHT.

One problem with this is that it threatens to make the claim that the chosen paradigmatically normative concept is itself normative a trivial truth. It is important to see why this is a problem and what the problem is. In one sense of “trivial”, it may simply be trivial that the chosen paradigmatic concept is normative: it may be *uncontroversial* and *obvious* that it is. But the point is that in another sense it is not trivial: it is a *substantive*, albeit perhaps obvious, claim that (for example) REASON is normative. Compare the discussion of analyticity. A well-known understanding of analyticity is as Frege-analyticity: a sentence is analytic in this sense if and only if it is a logical truth or can be transformed into one by substitution of synonyms for synonyms.⁸ But this characterization (whatever other problems it may face) runs into problems regarding the analyticity of logical truths. If this characterization captures what it is for a sentence to be analytic, then to say that logical truths are analytic is just to say that logical truths are logical truths. That is completely trivial. One might have thought that to say that a sentence is analytic has some significant metaphysical and epistemological implications. Return now to the claim that REASON is normative. One might have thought that this is to say something significant, albeit relatively uncontroversial, about REASON. But on the current proposal, it threatens to amount to nothing more than saying that REASON stands in the right sort of relation to itself.

Of course, and importantly, saying that *what it is* for something to be normative cannot be understood as suggested is not by itself to deny the weaker claim that *necessarily*, a concept C is positively normative just in case C stands in the right relation to some paradigmatic normative concept. That is a separate issue.

A third kind of suggestion is that a concept’s being normative is a matter of it playing the right kind of psychological role, whatever exactly it is, related to for example deliberation, motivation and action-guiding. Expressivists prominently defend views like this, but one can

⁸ Boghossian (1996), p. 366.

defend this kind of view without being an expressivist.⁹ I have used the label *normative role* for this kind of psychological role, and one might call the suggested view on what makes a concept *the normative role view*. This is the view I am inclined to favor.¹⁰ However, problematic questions can be raised regarding this view too. One question about this view, raised by Stephen Finlay, is whether this characterization of the normative is broad enough. Whatever in the end the fate of the normative role view, it is initially plausible when it comes to concepts like OUGHT and REASON.¹¹ (Again to stress, the words “ought” and “reason” have many different uses, but the view seems plausible regarding the concepts they express as they occur in contexts of special importance to metanormative inquiry.) But even someone agreeing that the normative role view is plausible in that case can wonder whether this generalizes appropriately to all normative concepts. One possible response to this in turn is to restrict the normative role view to a class of normative concepts where it seems especially promising. For example, one can restrict it to the authoritatively normative and to say that a different account must be given of the formally normative. Any move like this raises questions about what is the proper target of analysis. Do the authoritatively and the formally normative together constitute a class such that we should expect a unified account of what it takes to be a member of that class?

A normative role view is naturally associated with motivational judgment internalism, according to which there is some sort of a necessary connection between making normative judgments and being motivated in a corresponding way.¹² But a normative role view does not have to be developed in such a way as to commit to such internalism. Motivation need not enter into it.¹³

A different kind of suggestion regarding what characterizes normative concepts as such is that those concepts are associated with a distinctive kind of *phenomenology*.¹⁴ Important to keep in mind when assessing any such suggestion is the distinction between on the one hand the claim that the use of normative concepts in fact generally is associated with a specific kind of phenomenology and on the other hand the claim that this is what

⁹ The most clear and explicit discussion of this theme in an expressivist context is in Gibbard (2003). Some texts where this view is defended or else takes center stage are Eklund (2017), Wedgwood (2007, 2018), McPherson (2018), Williams (2018), Finlay (2019), and Dunaway (2020).

¹⁰ See Eklund (2017), ch. 4.

¹¹ The expression “ought” and “reason” arguably have different kinds of uses, and when in the text I speak of *the* concepts OUGHT and REASON, this should be understood to mean the concepts these expressions express in relevant uses. Needless to say, I am sliding over possible theoretical complications, and some of them may be significant.

¹² See Rosati (2016), section 3.2, for an overview.

¹³ See Plunkett (2020) for helpful discussions regarding the different kinds of things normative role can come to.

¹⁴ See Bedke (2019, 2020) and Werner (forthcoming).

distinguishes a normative concept as normative. It is the latter that the suggestion demands. And then the following sort of objection suggests itself. What if some community uses seemingly normative concepts in many ways like we do, including for action-guiding purposes, but their use of the concepts is accompanied by a very different (or perhaps no?) phenomenology. The present proposal would then class these concepts of theirs as non-normative. But is that not the wrong verdict?¹⁵

This concludes my brief survey on views on what makes a concept normative. I have not here tried to settle the issue. But again, my own preferred view is the normative role view.

4. Distinctions

Let me now return to the supposed distinction between formal and authoritative normativity. While there plausibly is a distinction there, it is hard to characterize what formal normativity and authoritative normativity are supposed to be. In a recent overview article on varieties of normativity, Derek Baker (2017) explains formal normativity as follows: “Formal normativity is the normativity displayed by any standard one can meet or fall short of. The rules of chess are formally normative, as are club rules, ancient honor codes, the law, and the standards of beauty employed by the Miss America pageant”.¹⁶ The examples should serve to convey what the idea is, and there plausibly is some sort of distinction of the kind at issue. But Baker’s account of formal normativity in terms of “standards” threatens to overgenerate. Even paradigmatically non-normative concepts and properties are associated with standards in a perfectly straightforward way. *Being over four feet wide* is associated with the standard of being over four feet wide. In his discussion of the formal normativity, Tristram McPherson (2009) says that “Chess is formally normative simply in virtue of its being possible to play an

¹⁵ Matthew Silverstein (2017) defends a view on which it is characteristic of normative concepts that they are about deliberation. He says: “The conclusion of ethical deliberation— a judgment about what I ought to do—is just a belief about the outcome of sound practical reasoning”.¹⁵ There are two ways to understand what Silverstein says. One possibility is that he means that facts about what an agent ought to do are identical to facts about sound practical reasoning. Another is that he also means to say something about the concept OUGHT. Compare: it is one thing to say that the *property* of being water is identical to the *property* of being H₂O, and another to say that the *concept* WATER is identical to the *concept* H₂O, so that to think the thought *there is water in the lake* is to think the same content as thinking *there is H₂O in the lake*. It is the latter claim that is most immediately relevant when the topic is normative concepts. In both cases, the former claim, about the world, is initially more plausible than the latter claim, about representation. It seems one can think WATER thought without being able to think H₂O-thoughts. And, returning to Silverstein, it seems that someone can think OUGHT thoughts without thinking thoughts, or even being able to think thoughts, that are directly about reasoning. A natural reply on Silverstein’s behalf is of course that what a thinker in fact thinks when employing a certain concept need not be transparent to the thinker. But even granting the general point about non-transparency, the question is whether we can be radically mistaken about the contents of our thoughts in the way we would seem to be given Silverstein’s view.

¹⁶ Baker (2017), p. 568.

incorrect chess move”.¹⁷ The general idea would be that something is formally normative by virtue of being associated with standards of correctness and incorrectness. This might be what underlies Baker’s characterization, and it gets around the counterexample I presented. Being over four feet wide, while associated with a standard in some sense (something meets it just in case it is over four feet wide), is not intuitively associated with a standard of *correctness*. Of course, McPherson’s appeal to correctness invites questions about what it is for something to be or not to be associated with a standard of correctness. Other characterizations of the distinction between formal normativity are problematic in other ways. Derek Parfit (2011) characterizes authoritative normativity as “reason-implying” and formal normativity as (merely) “rule-implying”.¹⁸ This seems no more successful, for it is unclear why the formally normative should not be regarded as being reason-implying. Even if the normativity of chess is merely formal, it seems that we can say that white has reason to move the queen, where this is in no way conditional upon a hypothesis about the player’s actual aims. (Stephen Finlay (2019) reflects, “A certain chess move might be a good move because it is likely to provoke one’s opponent into a rash response, which is a *reason* to make it, but doesn’t obviously imply any *rule*”.¹⁹)

A common feature of these characterizations of formal normativity they all seem in the first instance to be characterizations of what it is for *phenomena in the world* to be formally normative. The question of what makes a *concept* formally normative is different. Take McPherson’s characterization, appealing to correctness, as an example. He is talking about what makes chess, or a chess move, formally normative. If one wants to use McPherson’s suggestion and give an account of what makes concepts formally normative one needs to say something like, for example, that a concept is formally normative just in case competence with the concept involves taking what the concept applies to as having associated correctness conditions.

Baker’s and McPherson’s characterizations of formal normativity can also be proposed as characterizations of normativity full stop: they give a weak characterization satisfied by everything that is normative. A natural task for them, assuming the characterizations of the formally normative is on the right track, is to distinguish what is authoritatively normative from what is merely formally normative. McPherson presented the

¹⁷ McPherson (2011), p. 232.

¹⁸ Parfit (2011).

¹⁹ Finlay (2019), p. 45; my emphasis.

following proposal.²⁰ A concept is authoritatively normative by virtue of being a concept whose role it is to settle deliberation, such as for example deliberation over whether to let moral considerations override prudential ones or vice versa.²¹ I will not here get into the details of McPherson’s discussion, but will just note that – as McPherson himself notes – on his characterization moral concepts are not authoritatively normative. Whether or not this consequence is acceptable, it is certainly striking, for the normativity associated with morality does in an intuitive sense seem more authoritative than that associated with chess and etiquette. Baker takes a more radical tack. He discusses and rejects some different possible characterizations of authoritative normativity but despairs of the possibility of making sense of the notion.²² Parfit’s characterization is different being reason-implicating (the condition for being authoritatively normative) does not in any natural way imply being rule-implicating. The most natural way to turn Parfit’s characterization into a characterization of normativity generally is simply to present a disjunctive account: what is normative is what is either reason- or rule-implicating.

Let me now turn to the distinction between thin and thick concepts. I have discussed this at some length elsewhere so let me here just briefly indicate some of the problems regarding how to characterize this distinction.²³ Bernard Williams (1985), who introduced the terminology of “thick” and “thin” concepts, describes thick concepts as “world-guided”: their application “is determined by what the world is like”.²⁴ This is supposed to contrast with the more purely action-guide role of thin concepts. But is not also the application of the concept GOOD determined by what the world is like, and so world-guided? Allan Gibbard (1992) holds that a term stands for a thick concept if it praises or condemns an action as having a certain property.²⁵ One may be uneasy already about the talk of terms (as opposed to users thereof) praising or condemning actions. But take such talk on board. There is still a problem. There is as much reason to think that ‘good’ satisfies this condition as that an intuitively term like ‘courageous’ does. ‘Good’ stands for the property of being good. And is it not equally plausible that this term praises actions as having this property, just as ‘courageous’ praises actions as being courageous?²⁶

²⁰ See, e.g., McPherson (2018) and (2020).

²¹ See primarily McPherson (2018).

²² Baker (2017), pp. 577ff

²³ Eklund (2011).

²⁴ Williams (1985), p. 29.

²⁵ Gibbard (1992), 268f.

²⁶ In my (2011) I discuss, and reject, possible responses on Gibbard’s behalf to appeal specifically to *descriptive* properties.

In his seminal (2013) book on thick concepts, Pekka Väyrynen argues that thick terms – the expressions we would normally say express thick concepts – are not evaluative as a matter of meaning. They are used to convey something evaluative but do so through pragmatic means. On Väyrynen’s view there are not any thick normative concepts in the sense of thick concepts that are normative by virtue of aspects of their content.²⁷ One question about Väyrynen’s view is whether one can coherently maintain this view on thick concepts but resist it in the case of thin concepts.

5. The reference of normative concepts

Concepts, including normative concepts, often refer; they stand for things in the world. One may speak of this in different ways. One might speak of concepts as standing for properties (“the concept RIGHT stands for the property of being right”), or one may speak of concepts as standing for the things in the world that they are true of (“the concept RIGHT stands for the things that are right”). I will here help myself to property talk but I do not see that much or anything hinges on this.

How is it determined what a given normative concept stands for? Some theorists, most prominently in the expressivist tradition, might reject that question. They might deny that normative concepts stand for anything, or at least that normative concepts stand for anything in more than a “minimalist” sense, where “minimalist reference” is such that somehow all that can be said about reference is trivial (paradigmatically: the concept F refers to being F).

In the literature, naturalist realists, who hold that normative concepts stand for properties that are natural (in the sense of being, roughly, properties of the same general kind as those dealt with by the sciences) have tended to devote the most attention to the question of reference-determination, presumably because it is felt that it is *prima facie* puzzling how normative concepts could stand for such properties and naturalists have felt compelled to deal with this puzzle.²⁸ Perhaps unsurprisingly, the theories of reference-determination that they have come up with have amounted to applications to the case of normative concepts of theories that have been proposed as more general theories of reference-determination.²⁹

Cornell realists have defended causal theories according to which normative concepts stand

²⁷ For critical discussion of Väyrynen, see e.g. Roberts (2015) and Eklund (2017a).

²⁸ See Ridge (2019) for more on what “natural” means in the context. For a discussion of the issue of reference-determination as it arises for *non-naturalists*, see Werner (2020). In his work, Werner appeals centrally to the work of Imogen Dickie, specifically Dickie (2015) and (2016).

²⁹ In the main text I speak of the relevant theories of reference-determination as concerning *normative* terms and concepts generally. Proponents of the theories tend to speak as if they focus specifically on *moral* terms and concepts.

for what in the world their use is appropriately causally connected to.³⁰ Just as, somehow or other, “water” refers the stuff in the world that the use of the term is appropriately causally connected to, “good” refers to that property in the world that the use of the term is appropriately causally connected to. An apparent consequence is that otherwise identical speakers in principle can refer to different things, just by being in different causal environments. *Neo-descriptivists* have defended the view that what normative terms refer to is determined by associated folk theories (collections of claims that have the status as platitudes) analogously to how the reference of some scientific terms are sometimes held to be determined by scientific theories. The terms refer to whatever makes the associated theory true (or comes closest to doing so and close enough, where nothing perfectly fits the bill).³¹

A prominent argument against these theories of reference-determination is the so-called Moral Twin Earth argument, due to Terence Horgan and Mark Timmons.³² In brief, the Moral Twin Earth argument notes that it seems possible, on either of these theories, that the term “good” could play the same normative role for two communities but still have different reference as used by these communities – and then goes on to argue that this consequence is problematically counterintuitive. The Moral Twin Earth argument has generated extensive discussion and friends of the views criticized have of course sought to respond to it.³³

Let two concepts be *normative counterparts* if they have the same normative role but different reference. Then what the Moral Twin Earth argument says is that theories of reference-determination like those mentioned allow for the possibility of normative counterparts of GOOD, BAD, RIGHT and WRONG, and that this consequence is counterintuitive. These are two different claims. A first issue is whether it is possible for there to be normative counterparts of thin concepts, given a particular theory of how the reference of normative concepts is determined. A second issue is whether this consequence is unacceptable.

The causal and neo-descriptivist theories of how the reference of normative concepts is determined are theories that have the consequence that concepts can have normative counterparts. A theory according to which the reference of a normative concept is determined by the normative role associated with the concept rules out this possibility.

Assuming that it is possible for there to be normative counterparts of OUGHT that ascribe different properties, it seems that in some sense there is a question of which concept to

³⁰ See for example Boyd (1988) and Brink (1989).

³¹ See Jackson (1998).

³² See for example their (1992a), (1992b), and (2009).

³³ For some responses, see Copp (2000), Merli (2002), Dowell (2016), Dunaway and McPherson (2016), and Väyrynen (2018).

use. One may want to ask: ought I to use our concept OUGHT or a normative counterpart, OUGHT*, ascribing a different property? If I know that there exists this counterpart OUGHT*, then simply sticking to my original concept, OUGHT, seems arbitrary. At the same time it is hard to see how best to state what is supposed to be at issue when I ask which concept to use. It may well be that if I express the question in the way I just did, as

Ought I to use OUGHT or OUGHT*?

then the answer I arrive at is that I ought to use OUGHT – but what am I to do with that piece of information if when I express the question as

Ought* I to use OUGHT or OUGHT*?

then the answer is that I ought* to use OUGHT*? Issues like the ones I have just described are at the center stage of my own (2017).

In the above discussion of reference-determination, I have discussed matters as if one and the same theory of reference-determination is true of all normative concepts. But one may wonder what justifies such an assumption. First, among the normative concepts that we have are concepts of a variety of kinds. Why should we assume that these all work the same way? Second, even if it turned out to be the case that all of *our* normative concepts work the same way, the concepts that we have may well be only some of all the possible normative concepts that there are.

In my (2017), I used the possible existence of normative counterparts to present problems for what I call *ardent realism*. The intuitive thought behind ardent realism is that reality itself favors certain ways of acting. One might have thought that for this to be so it suffices that there are objective truths about what ought to be done, what it is right to do, what there is reason to do, etc. But if there *also* are objective truths about what ought* to be done, what it is right* to do, etc then the question seems to arise: which concepts, ours of the counterparts, track what reality favors? (Shamik Dasgupta (2017) and Justin Clarke-Doane (2020) have presented arguments in much the same spirit, even though are important differences between our arguments.)

I invented the label “ardent realism” for a reason. Normative realism, as it is often discussed, is simply the view that normative sentences are apt for truth, and some atomic sentences are mind-independently true. Such realism is perfectly compatible with the idea that

alongside mind-independent truths about what ought to be done, there simply are other mind-independent truths about what ought* to be done. She need not embrace any further idea about what reality favors. I believe many self-identified realists are in fact ardent realists but that does not immediately fall out of the realism as it tends to be characterized.

One natural thought when one faced with the question I raise for ardent realism is to take a leaf from metaphysics. In contemporary metaphysics, it is popular to hold that certain properties are more elite/natural/ fundamental/joint-carving—henceforth I will just use *elite*—than others.³⁴ The more elite properties are for example more explanatory, and are what account for objective similarities. Compare the property of being green with the property of being grue, where something is grue exactly if it is either green and observed before the year 2100 or blue and not so observed.³⁵ The property of being green seems more elite. Given the idea of eliteness, one might further suggest that the normative concepts that pick out the most elite properties are the ones that track what reality favors. But the general idea underlying the talk of what is elite is that the elite properties are the ones that correspond to objective similarities in the world, and what is fundamentally explanatory. And some normative concepts can track objective similarities and stand for something explanatory, for reasons independent of how well or not they discharge their normative functions.³⁶

Even if appeal to eliteness is, for reasons briefly indicated, unsuccessful as a response to the challenge that I have raised for ardent realism, it can in principle be relevant as a response to the Moral Twin Earth problem. The idea would be that a friend of a theory of the kind threatened by the Moral Twin Earth example appeal to the eliteness of the properties picked out by the normative concepts at issue. It is common to think that, somehow or other, the eliteness of a property makes it objectively more eligible to be referred to than something less elite—as it is often put, it is reference-magnetic. This reference-magnetism can be appealed to in an argument that both communities refer to the same properties after all.³⁷

In his (2018), Robbie Williams says

Some say that the concept *moral wrongness* has a distinctive referential stability. Environment, society and moral opinion may vary dramatically, and yet agents will succeed in thinking about a common subject matter, wrongness, so long as deliberation and sentiments are internally regulated in the right way. Schematically: there is a

³⁴ The classic references are Lewis (1983, 1984). See too Sider (2011).

³⁵ See Goodman (1955).

³⁶ See Eklund (2017), section 2.3, and Clarke-Doane (2020), p. 169f.

³⁷ See Dunaway and McPherson (2016).

property P, applying to actions, which our concept of moral wrongness picks out, and there is also a conceptual role R which our concept of moral wrongness plays. The *referential stability thesis* is the following: necessarily if an agent has a concept W that plays role R, then W denotes P.³⁸

It is clear from the context that the conceptual role Williams speaks of is what I here call normative role.

The thesis that MORAL WRONGNESS is referentially stable in the way outlined – what Williams calls the referential stability thesis – can be motivated by reflection on Moral Twin Earth arguments. Much of Williams’ discussion is then devoted to an explanation of the truth of this referential stability thesis. Generally, note that assuming that this referential stability is a datum, one can want to appeal to reference-magnetism to explain it. What Williams offers is one version of this, appealing to a particular kind of explanation of what reference-magnetism is and why it obtains.

The label “referential stability thesis”, and the way Williams discusses this thesis, might give the false impression that the claim is that MORAL WRONGNESS (to take Williams’ example) is generally speaking plainly more referentially stable than concepts typically are. But what is at issue is only stability regarding *some salient dimensions* – as Williams says, it seems that environment, society and moral opinion may vary dramatically while the reference stays the same. But along another dimension there can be significant instability, compatibly with the referential stability thesis as Williams characterizes it: already a minute difference between normative role R and normative role R* can yield that a concept that plays R* instead of R denotes a property distinct from P.

The so-called referential stability thesis plays a role also in my (2017) discussion. Without taking a stand on its overall plausibility, I discuss it as possible response to the challenge I described above. The suggestion would be that if the referential stability thesis is true then the challenge cannot be raised. (However, as I discuss, the challenge can be raised in other ways even if the thesis is true. For example, one might appeal to things like concepts associated with somewhat different normative roles.³⁹ This is relevant also to the Moral Twin

³⁸ Williams (2018), p. 41. The second italicization is mine.

³⁹ See Eklund (2017), ch. 3.

Earth argument. As presented, it assumes that the concept at issue have the same normative roles. But this assumption can perhaps be liberalized.)

It is of course a good question whether the referential stability thesis, or any similar thesis, is true. Both Williams and I discuss ways in which the thesis must be qualified. Much of the discussion in Billy Dunaway (2020) is centered on problem cases for the referential stability thesis.⁴⁰ Dunaway thinks that the true thesis in the vicinity of the referential stability thesis is instead the following, somewhat weaker thesis:

Robust Disagreement. Any two communities that use a primitive term with the same moral or normative role, and differ at most in which substantive theory they follow in applying the relevant term, are thereby capable of having a substantive disagreement with each other with that term.⁴¹

6. Normative concepts and conceptual engineering

In recent philosophy, *conceptual engineering* has become a hot topic.⁴² In brief, whereas conceptual analysis as traditionally practiced seeks to understand what certain concepts we actually have are like, conceptual engineering is focused on also on what concepts there can be, and specifically on what concepts might in some sense be better than the concepts we actually employ.

Conceptual engineering has in effect come up in the discussion already, even if I have not introduced the label until now. My own challenge to ardent realism centrally involves appeal to possible alternative normative concepts and hence relates to conceptual engineering.

Another inroad into conceptual engineering of normative concepts comes via consideration of specific problems with actual normative concepts we have. Here is one illustration. As mentioned already earlier, some normative concepts seem somehow or other to be normatively *objectionable*. Arguably, many concepts pertaining to sexual morality, such as LEWD, CHASTE, PROMISCUOUS, and (SEXUALLY) PERVERTED are in this category. These concepts ought not to be used for evaluation and prescription: they embody mistaken normative assumptions. Arguably, we should abandon them and not employ them—at least not for these purposes. As far as I am concerned, such abandonment can count as one form of

⁴⁰ Dunaway (2020), ch. 2.

⁴¹ Dunaway (2020), p. 59.

⁴² For some prominent discussions, see Cappelen (2018) and the essays in Burgess, Cappelen and Plunkett (2020).

conceptual engineering, even if it is not a matter of revising or replacing concepts. But one can also ask whether there are revised versions of these concepts that are not objectionable. For example, while our actual concept CHASTE seems somehow to rely on the normative assumption that sexual abstinence is in itself a good thing, a revised version might instead be such that someone counts as chaste exactly if they conduct their sex life with dignity and self-respect.

More generally, at least from any broadly realist standpoint the following seems an obvious point to make. When we think of normative concepts, we naturally think of the specific normative concepts that we have and employ. But more generally, what kinds of normative concepts can there be? Just as the concepts that (say) present and past physics employs are not all of the physics concepts there can be, so too the concepts that we do employ and have employed in normative thought need to be all the normative concepts that there can be. There may be different normative properties out there, and in order to ascribe them we need concepts we do not currently employ.

A different, more radical kind of conceptual engineering of normative concepts is the following. All sorts of different theories of normative language and thought have been presented and defended – non-naturalist and naturalist cognitivist theories, prescriptivism, expressivism, and various forms of hybrid cognitivism and hybrid expressivism.

Briefly and roughly, cognitivist theories hold that normative language and thought functions like paradigmatic descriptive thought does, in terms of representing the world. They differ over whether they take normative language and thought to concern natural properties, as naturalists hold, or non-natural properties, as non-naturalists hold. Prescriptivism holds that normative language and thought is primary for prescribing. (It is more naturally pitched as a theory of language.) Expressivists tend to hold that the normative fundamentally has to do with non-cognitive attitudes. As the names indicate, hybrid theories hold that normative concepts have a dual nature and have both cognitivist and expressivist components.⁴³

These theories are all theories of how actual normative thought and language work. But regardless of which theory gets this right, there is a separate question of whether some community's normative thought and language could work in such a way that another theory is true of them. And if different theories can be true of the normative thought and language of

⁴³ For helpful overviews of these different –isms which get into some more detail, see Lutz and Lenman (2018), Ridge (2019), and Toppinen (2017). These –isms are traditionally proposed and understood as theories of the *moral*.

different communities, might there be reasons for thinking that some other form of normative thought and language is better than the actual one?

For example, even if some expressivist theory is true of our normative thought and language, maybe some possible community could use normative language of which cognitivism is true in order to state facts about the world: there are properties and facts, of the kind the realist cognitivist believes in, for their normative expressions to refer to and state. Perhaps we are then missing out on something and ought instead to think and speak like them?

What I am mentioning here is something more radical than the kinds of conceptual engineering earlier mentioned, and than the examples of conceptual engineering typically discussed in the literature. It is not merely a matter of replacing a concept standing for one property by a concept standing for another property, but it is a matter of replacing concepts of a certain kind by concepts of very different kinds.

Needless to say, some theorists will hold that, for example, there simply are no non-natural properties for any concepts to pick out, so a language which is non-naturalist in the sense that its normative predicates are semantically constrained to pick out non-natural properties if they pick out any properties at all is a language whose normative predicates fail to pick out any properties. But whether or not it is so is independent of what happens to be true of our language. Even assuming a compelling argument that our language works as the expressivist or naturalist realist says it does, there could be non-natural properties for the normative predicates of a suitable non-naturalist language to pick out.

Someone with expressivist leanings may well be inclined to say that even if there are such non-natural properties for normative predicates and concepts to pick out, we are not missing out on anything by not using such normative predicates and concepts but are doing just fine: an expressivist language does for us everything one could need normative language for. That may be so. But the claim that it is seems to be a *more* central contention about the nature of normativity than any claim merely about how our actual normative language and thought work.

One version of the expressivist response is to say that predicates and concepts that function in a descriptivist way would not be normative at all. In order for this complaint to be of philosophical interest it would need to say that any descriptivist predicates and concepts would be so different from actual normative predicates and concepts that there can be no real similarity in function and role. Any weaker version of the complaint would reasonably be met with the response that while on a strict understanding of “normative” maybe the descriptivist

concepts would not be normative, they are still sufficiently normative-ish (they may be “normative*”) to serve as replacements.

There are different kinds of missing out. Here is one distinction, implicit in the above but worth stressing. We are *normatively* missing out if some other possible language serves some worthwhile normative – action-guiding – purposes not served by our actual language. We are *descriptively* missing out if there is some property – such that there is some reasonable purpose served by picking out that property – such that some other language contains means to pick out that property but our language does not. The kind of expressivist just envisaged can allow that we are descriptively missing out insofar as we lack the means to talk about non-natural properties. In fact, it may be easy to descriptively miss out: for it may be that for every property there is a reasonable purpose served by picking out that property – the purpose of saying true things about the property in question. What is important for the envisaged expressivist is only that we are not normatively missing out. She can allow that it would be useful for us to have some means of picking out the properties in question. What she would disagree with is the claim that any legitimate normative purposes are so served. It would be natural for her to say that it is immaterial (at best) that normative language is used to pick out these properties.

Conversely, someone can subscribe to a cognitivist understanding of actual normative language and thought while yet thinking that an expressivist theory of normative language and thought describes a workable practice – and that nothing would be lost (and maybe even something gained) by instead engaging in expressivist normative thought and talk.

Within the class of representational systems described by cognitivist theories, compare possible representational systems where normative concepts stand for natural properties and possible systems where such concepts stand for non-natural properties. If there are systems of both kinds, one can then ask what sort of representational system is better for normative purposes. Maybe as a matter of fact representational systems where the normative concepts ascribe non-natural properties are better for normative thought and talk than a system where they ascribe natural properties. But this would have to be argued for. And even if as a matter of fact our normative concepts ascribe non-natural properties, that does not mean that there is some extra value in ascribing such properties.

One central topic in debates over naturalism and non-naturalism is whether necessarily coextensive properties are identical. Some prominent naturalist arguments assume that it is

so.⁴⁴ But suppose, with the non-naturalist in this debate, that necessarily coextensive properties are not identical. There is still a live question regarding the extent to which it matters which of two necessarily coextensive properties a given normative concept picks out. It can be that even if there are non-natural properties for normative predicates to pick out, and even if those are the properties that our normative predicates pick out, a language whose predicates instead pick out the necessarily equivalent natural properties is just as good for normative purposes. (And hence these natural properties are equally deserving of the title “normative”.)

And from an expressivist standpoint one can compare different expressivist languages – for example, ones where normative predicates express certain attitudes and ones where these predicates express different attitudes – and ask which kind of language is more apt for normative purposes.

7. Open questions

Let me close by just briefly stressing the underexplored questions I have devoted some time to discussing here. There is much more to say about all of them, and I hope philosophers will undertake this task in future work. One question concerns what it is for a concept to be normative. There has been surprisingly little systematic work devoted to this. My own favored view is what I here call the normative role view. But this view faces unanswered question. Other underexplored questions concern what determines the reference of normative concepts (on the assumption that these concepts refer at all), and what kind of thesis concerning referential stability of normative concepts is true (if any). Lastly, there are the broad questions, highlighted in the last few sections, going beyond questions about *actual* normative concepts and instead being about what sorts of *possible* normative concepts there are.

REFERENCES

Baker, Derek: 2017, “The Varieties of Normativity”, in Tristram McPherson and David Plunkett (eds.), *The Routledge Handbook of Metaethics*, Routledge.

⁴⁴ One example is Frank Jackson’s supervenience argument, presented in his (1998). For discussion, see McPherson (2019), section 3.2.

- Bedke, Matthew: 2019, “Choosing Normative Concepts (Review)”, *Philosophical Review*, 128: 121–126.
- Bedke, Matthew: 2020, “What Normativity Cannot Be”, *Journal of Ethics and Social Philosophy* 18.
- Boghossian, Paul: 1996, “Analyticity Reconsidered”, *Noûs* 30: 360-91.
- Boyd, Richard: 1988, “How to be a Moral Realist”, in Geoff Sayre-McCord (ed.), *Essays on Moral Realism*, Cornell University Press, Ithaca, New York, pp. 181–228.
- Brink, David: 1989, *Moral Realism and the Foundations of Ethics*, Cambridge University Press, Cambridge.
- Burgess, Alexis, Herman Cappelen, and David Plunkett (eds.): 2020, *Conceptual Engineering and Conceptual Ethics*, Oxford University Press.
- Cappelen, Herman: 2018, *Fixing Language*, Oxford University Press.
- Copp, David: 2000, “Milk, Honey and the Good Life on Moral Twin Earth”, *Synthese* 124: 113-137.
- Dancy, Jonathan: 2006, “Nonnaturalism”, in David Copp (ed.), *The Oxford Handbook of Ethical Theory*, Oxford University Press, Oxford, pp. 122–45.
- Dowell, Janice: 2016, “The Metaethical Insignificance of Moral Twin Earth”, *Oxford Studies in Metaethics* 11: 1–27.
- Dunaway, Billy: 2020, *Reality and Morality*, Oxford University Press.
- Dunaway, Billy and Tristram McPherson: 2016, “Reference Magnetism as a Solution to the Moral Twin Earth Problem”, *Ergo* 3: 639–79.
- Eklund, Matti: 2011, “What are Thick Concepts?”, *Canadian Journal of Philosophy* 41:25-49.
- Eklund, Matti: 2017, *Choosing Normative Concepts*, Oxford University Press.
- Eklund, Matti: 2017a, “Thickness and Evaluation”, *Journal of Moral Philosophy* 14: 89-104.
- Finlay, Stephen: 2019, “Defining Normativity”, in David Plunkett, Scott Shapiro and Kevin Toh (eds.), *Dimensions of Normativity: New Essays on Metaethics and Jurisprudence*. Oxford University Press.
- FitzPatrick, William: 2011, “Ethical Non-Naturalism and Normative Properties”, in Michael Brady (ed.), *New Waves in Metaethics*, Palgrave Macmillan, London, pp. 7–35.
- Gibbard, Allan: 1992, “Morality and Thick Concepts (I) – Thick Concepts and Warrant for Feelings”, *Proceedings of the Aristotelian Society*, suppl. vol. 66: 267–83.
- Gibbard, Allan: 2003, *Thinking How to Live*, Harvard University Press, Cambridge, Massachusetts.

- Goodman, Nelson: 1955, *Fact, Fiction, and Forecast*, Harvard University Press, Cambridge, Massachusetts.
- Horgan, Terence and Mark Timmons: 1992a, “Troubles for New Wave Moral Semantics: The ‘Open Question Argument’ Revived”, *Philosophical Papers* 21: 153–75.
- Horgan, Terence and Mark Timmons: 1992b, “Troubles on Moral Twin Earth: Moral Queerness Revived”, *Synthese* 92: 221–60.
- Horgan, Terence and Mark Timmons: 2009, “Analytical Moral Functionalism Meets Moral Twin Earth”, in Ian Ravenscroft (ed.), *Minds, Ethics and Conditionals*, Oxford University Press, Oxford, pp. 221–37.
- Jackson, Frank: 1998, *From Metaphysics to Ethics*, Clarendon Press, Oxford.
- Lewis, David: 1983, “New Work for a Theory of Universals”, *Australasian Journal of Philosophy* 61: 343–77.
- Lewis, David: 1984, “Putnam’s Paradox”, *Australasian Journal of Philosophy* 62: 221–36.
- Lutz, Matthew and James Lenman: 2018, “Moral Naturalism”, in Ed Zalta (ed.), *Stanford Encyclopedia of Philosophy*. Available at <https://plato.stanford.edu/entries/naturalism-moral/>.
- McPherson, Tristram: 2011, “Against Quietist Normative Realism”, *Philosophical Studies* 154: 223–40.
- McPherson, Tristram: 2018, “Authoritatively Normative Concepts”, in Russ Shafer-Landau (ed.), *Oxford Studies in Metaethics*, vol. 13. Oxford University Press.
- McPherson, Tristram: 2019, “Supervenience in Ethics”, in Ed Zalta (ed.), *Stanford Encyclopedia of Philosophy*. Available at <https://plato.stanford.edu/entries/supervenience-ethics/>.
- McPherson, Tristram: 2020, “Ardent Realism Without Referential Normativity”, *Inquiry* 63: 489–508.
- McPherson, Tristram and David Plunkett: 2020, “Conceptual Ethics and the Methodology of Normative Inquiry”, in Burgess, Cappelen and Plunkett (2020).
- Merli, David: 2002, “Return to Moral Twin Earth”, *Canadian Journal of Philosophy* 32: 207–40.
- Moore, G.E.: 1903, *Principia Ethica*, Cambridge University Press, Cambridge.
- Parfit, Derek: 2011, *On What Matters, Volume 2*, Oxford University Press.
- Plunkett, David: 2020, “Normative Roles, Conceptual Variance, and Ardent Realism about Normativity”, *Inquiry* 63: 509–534.
- Ridge, Michael: 2019, “Moral Non-Naturalism”, in Ed Zalta (ed.), *Stanford Encyclopedia of Philosophy*. Available at <https://plato.stanford.edu/entries/moral-non-naturalism/>.

- Roberts, Debbie: 2013a, “It’s Evaluation, Only Thicker”, in Simon Kirchin (ed.), *Thick Concepts*, Oxford University Press, Oxford, pp. 78–96.
- Roberts, Debbie: 2015, “Review of Pekka Väyrynen, *The Lewd, the Rude and the Nasty*”, *Ethics* 125: 910–5.
- Rosati, Connie: 2016, “Moral Motivation”, in Ed Zalta (ed.), *Stanford Encyclopedia of Philosophy*. Available at <http://plato.stanford.edu/entries/moral-motivation/>.
- Silverstein, Matthew: 2017, “Ethics and Practical Reasoning”, *Ethics* 127: 353-82.
- Sturgeon, Nicholas: 2003, “Moore on Ethical Naturalism”, *Ethics* 113: 528-56.
- Toppinen, Teemu: 2017, “Hybrid Accounts of Ethical Thought and Talk”, in Tristram McPherson and David Plunkett (eds.), *Routledge Handbook of Metaethics*, Routledge, pp. 243-59.
- Väyrynen, Pekka: 2013, *The Lewd, The Rude, and the Nasty*, Oxford University Press, Oxford.
- Väyrynen, Pekka: 2018, “A Simple Escape from Moral Twin Earth”, *Thought* 7: 109-118.
- Wedgwood, Ralph: 2007, *The Nature of Normativity*, Clarendon Press, Oxford.
- Wedgwood, Ralph: 2018, “The Unity of Normativity”, in Daniel Star (ed.), *The Oxford Handbook of Reasons and Normativity*, Oxford University Press, pp. 23-45.
- Werner, Preston: 2020, “Getting a Moral Thing Into a Thought: Metasemantics for Non-Naturalists”, in In Russ Shafer-Landau (ed.), *Oxford Studies in Metaethics*, vol. 15, Oxford University Press. pp. 140-169.
- Werner, Preston: forthcoming, “Normative Concepts and the Return to Eden”, *Philosophical Studies*.
- Williams, J.R.G.: 2018, “Normative Reference Magnets”, *Philosophical Review* 127: 41-71.