

## Normative Ethics and Metaethics

For the *Routledge Handbook to Metaethics*, edited by David Plunkett and Tristram McPherson

By its very name, the field of metaethics is contrasted with another field of inquiry, that of normative ethics. Metaethics concerns questions *about* normative inquiry, rather than questions *within* normative inquiry. Whereas normative ethics concerns questions like what is good or bad, what we must or ought to do, and why, metaethics is said to be concerned with questions like what it means to say that something is ‘good’ or ‘bad’, whether such claims correspond to facts about the world, and how we know or manage to talk or think about such facts, if there are any. Metaethics may also be contrasted with normative ethics as philosophy of science is contrasted with science. Metaethicists, it is sometimes said, no more make claims within normative ethics than philosophers of science make claims within science, or pundits are engaged in playing football:

What is metaethics anyway? One useful way of answering this question is by contrasting metaethics with applied ethics and normative ethics. Consider an analogy that will illustrate the contrast: imagine ethics as football. We can equate different things associated with football with the different disciplines of ethics. . . there is the referee, who helps interpret the rules that the players are following. The referee can be thought of as the normative ethicist. . . Finally, there is the football analyst or pundit, who does not kick a ball or interpret the rules for the players but tries to understand and comment on what is going on in the game itself. This is like the metaethicist, who asks questions about the very practice of ethics, some of which we shall consider below. (Fisher [2011, 1-2])

On this conception of the distinction between normative ethics and metaethics, the divide between normative ethics and metaethics runs two ways. By its very nature, metaethics is concerned with questions that cross-cut the concerns of normative ethics, and so no answer to any metaethical question commits you in any way to any answer to a normative ethical question. Similarly, no answer to any normative ethical question commits you to any answer to a metaethical question. As we’ll see in what follows, this conception of the relationship between normative ethics and metaethics is mistaken. *Some* answers to *some* metaethical questions *do* carry commitments within normative ethics, and likewise, some answers to some normative ethical questions do carry commitments within metaethics. Indeed, as we’ll see, there is a very interesting set of views on which metaethics and normative ethics greatly overlap. Yet it is also a mistake to think that there

is no real distinction between normative ethics and metaethics. As we will see, there are, in fact, *some* answers to some very important and central metaethical questions that indeed do not carry any commitments for the answers to any normative ethical questions. Moreover, among the most prominent such answers, are the ones espoused by the philosophers who defended the sharp distinction between metaethics and normative ethics. So it is not hard to understand where the traditional conception of the distinction between normative ethics and metaethics could have come from.

I will *not* attempt, here, to provide a set of necessary and sufficient conditions on what counts as metaethics and what counts as normative ethics. Indeed, it is a consequence of my arguments that there may be no completely satisfactory way of drawing such a distinction, and I don't think that it is important to be able to draw one. Instead, I will rely in all of what follows on the classification of easy paradigms. Views which make claims about the nature of moral language, thought, epistemology, or reality, including noncognitivism, moral error theory, contextualism, relativism, and reductive realism, are all paradigmatically metaethical. Normative ethical theory, I will assume, paradigmatically includes both particular and highly general claims about what is good, best, right, wrong, or apt, as well as attempts to explain why. I will be arguing that many, but not all, paradigmatic metaethical views carry straightforward commitments for paradigmatic normative ethical questions, and that at least one kind of paradigmatically normative ethical view carries straightforward commitments for metaethics.

## **I      Traditional Noncognitivism**

One important and central question in metaethics concerns the nature of the meaning of moral words like 'good', 'ought', and 'wrong'. One important and central answer to this question is that we can get the most illuminating gloss on what gives these words their meanings, without either using these terms or applying predicates to them like 'true' or 'satisfies', which license inferences to sentences that use these terms. Instead, we understand what they mean by saying what they can be used to *do* – to praise or condemn, or to express attitudes of praise or condemnation. Answers which take roughly this form belong to the noncognitivist tradition in metaethics, and can be and have been developed in many different ways.

Noncognitivist accounts of the meaning of moral words are typically held by their proponents to be independent of commitments in normative ethics. To say what a moral word is used to do is not to use that word to do anything in particular, and so it is compatible not only with using the word 'wrong' to apply to, for example, stealing, and also with using the words 'not wrong' to apply to stealing, but with using the word 'wrong' in neither of these ways. Moreover, there are no valid inferences from sentences like 'the word 'wrong'

is used to condemn' to sentences that use the word 'wrong'. And this contrasts with more familiar, cognitivist or descriptivist, approaches to meaning, on which we say what the word 'wrong' means using sentences like 'the word 'wrong' is satisfied by all and only actions that fail to maximize happiness'. Semantic descent licenses us, from this premise, to infer, 'an action is wrong just in case it fails to maximize happiness'. Because noncognitivist theories do not rely on words like 'true', 'false', or 'satisfies' that license such inferences, they carry no similar commitments.

It is important to qualify this point, however. Some views which can for some purposes be classified as 'noncognitivist' are committed to answers to questions within normative ethics. Some noncognitivist views – sometimes called *hybrid* views, such as those developed in or discussed by Copp [2001], Boisvert [2008], Schroeder [2009], and Hay [2011] – model moral claims by analogy to epithets or slurs. According to these views, we cannot understand the meaning of moral words without understanding that they are used to endorse or to condemn, and that is what makes them count as 'noncognitivist'. But according to these hybrid forms of noncognitivism, it is not enough to understand the meaning of moral words, to know that they are used to endorse or to condemn; we must also understand what they are *about*, and hence what they are endorsing or condemning.

The relationship between normative ethics and metaethics is more complex for hybrid theorists. Just as you do not know the meaning of the word 'kraut' until you know that it is a slur for Germans, the hybrid theorist has not told us the meaning of the word 'wrong' until she has told us what it is used to condemn. A theorist can tell us that 'kraut' is a slur for Germans without taking a stand on whether anyone is a kraut; for if 'kraut' is not one of her words, because she does not have contempt for Germans, she may refuse to use it. So it is possible to say what the word 'kraut' means without taking a stand on who is a kraut. But unless the word 'wrong' is not one of her words in the very same sense, the hybrid noncognitivist cannot refuse to take a stand on which actions are wrong. Once she has told us what the word 'wrong' is used to condemn, she must accept that actions with those features are wrong, on pain of not being able to use the word 'wrong' at all. So her view carries commitments for normative ethics.<sup>1</sup>

---

<sup>1</sup> These remarks do not apply to a close relative of hybrid expressivism, *relational* expressivism. See Schroeder [2013] and Toppinen [2013] for discussion. Relational expressivism is also endorsed by Ridge [2014].

## 2 Error Theory

Another traditional and central question of metaethics concerns whether there are any moral truths at all. An important skeptical answer to this question is that there are not. Views that maintain as such are known as versions of the moral *error theory*.

Error theoretic views can take different forms. Some claim not only that there are no moral truths (as in Finlay [2014, IO-II]), but that moral claims are *false* (Mackie [1977, 35], Loeb [2007, 471], Olson [2014, 8]). This claim requires refinement. Compare the properties of being permissible and of being required. On a standard conception, these properties are *duals* – an action is permissible just in case its negation is not required. So if it is false that A is permissible, then A is not permissible, and hence  $\sim A$  is required – so it is true that  $\sim A$  is required. Conversely, if it is false that  $\sim A$  is required, then  $\sim A$  is not required, and so A is permissible – so it is true that A is permissible. If *permissible* and *required* are duals, therefore, then some very simple reasoning – which appeals only to very minimal assumptions about truth and falsity – leads to the conclusion that it is incoherent to say that *all* moral claims are false.

The simplest answer to this puzzle, leading to one intelligible form of error theory, is to take sides.<sup>2</sup> The natural side to take is to say that nothing is *required*. In fact, some of the standard motivations for the moral error theory are particularly naturally construed as arguments for precisely this thesis. For example, it is sometimes argued that nothing can be wrong unless there is a categorical reason not to do it, but that since there are no categorical reasons, nothing can be wrong. This is an argument that nothing is required (since an action is required just in case its negation is not wrong), but not an argument that nothing is permissible. On the contrary, given the duality of *permissible* and *required*, it is an argument that *everything* is permissible.

It should be clear that an error theory which takes this form is deeply committed to claims within normative ethics, and not only because ‘nothing is required’ sounds like a ‘first-order’ normative claim. Most ambitious explanatory normative ethical theories aspire to endorse biconditionals of the form, ‘for all  $x$ ,  $x$  is wrong just in case  $x$  is F’, where F is a non-trivial condition. The naïve error theory that I have just described is inconsistent with every such view, and so it carries commitments within normative ethics in a very deep sense.

However, there are other ways of developing the moral error theory. A different response to the problem of the duality of *permissible* and *required* is to hold that both ‘permissible’ and ‘required’ share a mistaken presupposition. On this view, ‘stealing is permissible’ says something like ‘stealing is permitted by

---

<sup>2</sup> Loeb [2007, 471], for example, restricts his characterization of error theory to the falsity of “positive” claims. Olson [2014, 14-15] says that all moral claims are false, but he defines “moral claims” so that ‘stealing is not wrong’ does not count.

code C', but carries some problematic presupposition about code C, for example that C is the code of conduct preferred by God. If no code of conduct is preferred by God, therefore, claims of the form 'stealing is permissible' will all carry a false presupposition. It is an important property of presuppositions that they *project*. This means that if 'stealing is permissible' and 'not stealing is permissible' both carry the presupposition that C is the code of conduct preferred by God, then so do 'stealing is not permissible' and 'not stealing is not permissible'. But by the duality of 'permissible' and 'required', these are just the claims that not stealing is required and that stealing is required. So if claims about permissibility carry a presupposition, claims about requirement carry the very same presupposition.<sup>3</sup>

The error theorist who holds that 'permissible' and 'required' carry a problematic presupposition will go one of two ways, depending on what she thinks it is most apt to say about claims that carry false presuppositions. She may say that both 'stealing is permissible' and 'stealing is not permissible' are *false*, because they carry the false presupposition that (in our example), C is the code of conduct preferred by God. Or she may say that neither is true nor false, because of their false presupposition. If she takes the latter course, then she will say that no moral claim is true, but she will not say that all moral claims are false. Instead, she will say that no moral claim is false, either. If she takes the former course, then she will say that all moral claims – at least, all claims about permissibility and requirement – are false. But this will not lead her into trouble with the problem of duality, because that problem started, at its first step, with the inference from, 'it is false that stealing is permissible' to 'stealing is not permissible'. But if we classify claims with false presuppositions as false, then this is not a valid inference – and indeed, it breaks down precisely when 'P' carries a false presupposition, and according to this form of the error theory, this is precisely such a case. So this form of the error theory evades the problem of duality, as well.

The forms of error theory that I have surveyed here are not the only ways of responding to the problem of duality. Another response to the problem of duality is to reject the premise that 'permissible' and 'required' are duals.<sup>4</sup> This can be motivated by the observation that 'permissible' has semantic structure – it is composed out of the parts 'permit' and 'ible'. And like other predicates of this structure, including 'laughable', 'laudable', and 'loveable', this makes 'permissible' plausibly analyzed as meaning something like 'aptly permitted' or 'appropriately permitted'. If we understand 'required' analogously, as 'aptly required', then there will be an alternative to stealing being either aptly permitted or aptly required – it may be that it

---

<sup>3</sup> Caleb Perl [ms] argues that all deontic claims carry presuppositions without arguing for the error theory. His idea is that this thesis solves multiple problems about the semantics of deontic vocabulary and in moral epistemology.

<sup>4</sup> Olson [2014, 14] takes this route, but not the following explanation. In his terminology, he denies that 'stealing is not wrong' entails 'stealing is permissible'. According to Olson, 'permissible' is a generalized conversational implicature of 'wrong', which is what leads us to assume that there is an entailment.

is neither aptly permitted nor aptly required, perhaps because nothing is apt. On this view, 'impermissible' means 'aptly impermitted', rather than 'not aptly permitted'.

Error theories which reject the duality of 'permissible' and 'required' in this way, like the naïve error theory we considered earlier, carry commitments in normative ethics. Their reason for claiming that moral claims like 'stealing is permissible' and 'stealing is impermissible' are both false, is that a third alternative is true – that it is neither apt to permit stealing nor to require not stealing. But this is a substantive claim about what is apt. Given that 'permissible' really does mean 'aptly permitted', as this view claims, it follows that every interesting substantive theory in normative ethics is committed to a non-trivial condition on whether something is apt. But this form of the error theory is inconsistent with every such condition on aptness. So it is incompatible with every substantive interesting theory in normative ethics.

In contrast, the forms of error theory which appeal to a problematic shared presupposition of moral words do not necessarily carry commitments for normative ethics. Of course, if the false shared presupposition is itself a moral claim – such as that code C is the one that it is best for everyone to follow – then this view has not avoided commitments in normative ethics, because it is committed to the view that the presupposition is false. So such error theories, at best, would avoid commitments about what is *wrong*, but would not avoid commitments about what is *best*. But if the presupposition is *not* itself a moral claim – for example, if it is the claim that code C is the code of conduct preferred by God – then the error theorist can claim that this presupposition is false without making any substantive moral claims. The advocate of presupposition who claims that all moral claims are truth-valueless does not make any claims from which any normative ethical claim can be deduced, because on her view 'it is not true that P' does not entail 'it is false that P', and hence does not entail ' $\sim$ P'. Similarly, the advocate of presupposition who claims that all moral claims are false does not make any claims from which any normative ethical claim can be deduced, because on her view 'it is false that P' does not entail ' $\sim$ P'.

Moreover, the failures of these entailments, on each view, are not *ad hoc* additions to each view, but follow immediately from the core claims that they make about the nature of moral language. Once we accept that moral claims carry any presupposition at all, we will get the failure of one of these entailments, and which one fails will be determined by our view about whether claims with false presuppositions are false, or merely lack truth values (see Strawson [1950]).

There is a sense, of course, in which even presupposition-based error theories preclude certain commitments within normative ethical theory. Endorsing such a theory, after all, is like responding to the question, 'when did you stop beating your wife?' by pointing out that you have never beat your wife. It would not make a lot of sense to point this out and then to go on to suggest that you stopped last week. Nor would

it make a lot of sense for the presupposition-based error theorist to go on to make claims about what is permissible or impermissible. That is why error theories constitute a threat to normative ethics. But in contrast to the naïve error theory and the duality-denying error theory, presupposition-based error theories *may* – depending on the presupposition which they attribute and deny – pose such a threat without carrying any commitments for the answers to any normative ethical questions whatsoever. Such views have consequences *for* normative ethical theory, but they do not have normative ethical consequences.

### 3 Contextualism/Relativism

As we've seen, traditional, non-hybrid, noncognitivist views appear to make claims that are orthogonal to normative ethics. In contrast, whether the moral error theory is committed or not to claims within normative ethics depends very much on how the error theory is developed, but there are *some* intelligible versions of the error theory which clearly carry no commitments to claims within normative ethics. Contrasting sharply with such views, are forms of *contextualism* and *relativism*.

According to contextualist metaethical theories, moral words may be used to make different claims in different contexts of utterance. Actually, even views that are not interestingly contextualist can accept this much. On a plausible and well-supported view about modal terms like 'must', 'may', and 'ought', for example, they are not ambiguous between distinct moral and non-moral senses, just because they can be used to make moral claims in some contexts, as in 'Jack must not steal – it's wrong', and also used to make clearly non-moral claims in other contexts, as in 'Jack must have stolen the cookie – no one else could reach the shelf'. Rather, such terms are context-dependent, and allow for both moral and non-moral claims to be made, given a single, contextually flexible, meaning. All interesting metaethical views can take this insight about modal words on board, but *interestingly* contextualist metaethical views such as that developed by Finlay [2014] maintain that even once we have fixed the context so that a moral reading of 'must' is required, different speakers or different conversations may fix on interestingly different claims as expressed by the very same sentence.

The bare claim that one and the same sentence – for example, 'stealing is wrong' – may express different propositions in the mouths of different speakers or in different conversational settings does not, in and of itself, carry any commitments for normative ethics. But no contextualist view is complete without telling us *how* context affects what claim can be made using such a sentence. The general form that such theories take, is to tell us, for an arbitrary context C, under what conditions the proposition expressed by 'stealing is wrong' in C is true.

Views of this structure do not avoid commitments in normative ethics. On the contrary, once we know which proposition is expressed by ‘stealing is wrong’ relative to an arbitrary context C, and know which context C we occupy, we can determine the conditions under which ‘stealing is wrong’ expresses a truth in our own context. In knowing which context we are in, we know a claim of the form, ‘C is our context’. And in knowing the consequences of our theory, we know a claim of the form, ‘ ‘stealing is wrong’ expresses a truth in C just in case S’, where ‘S’ is some sentence specifying the truth-conditions of ‘stealing is wrong’ in context C. And from these two claims, we can infer, ‘ ‘stealing is wrong’ expresses a truth in our context just in case S’. But no matter what context we occupy, claims of the form ‘ ‘P’ expresses a truth in my context just in case P’ express a truth in that context. So we are also committed to accepting the claim, ‘ ‘stealing is wrong’ expresses a truth in our context just in case stealing is wrong’. But now putting these two claims together, we can infer, ‘stealing is wrong just in case S’. But this is a normative ethical claim – indeed, it is a highly general one, asserting a condition of the highest possible generality for when stealing is wrong. In general, all that it takes to be committed to such claims, once we endorse a complete contextualist theory, is to know what context we are in.

Relativist metaethical theories are usefully contrasted with contextualist theories for most purposes, but are committed to normative ethical consequences in a very similar way. In general, whereas contextualists say that one and the same moral sentence may make different claims in different contexts of utterance, even once we have narrowed our attention to a particular moral use of the words in question, relativism such as that developed by MacFarlane [2014] and discussed in Brogaard [2008], as applied to metaethics, says that moral sentences make the *same* claim across such contexts, but that such claims are not absolutely true or false, but only relative to some perspective.

As with contextualism, the bare form of relativism does not carry any substantive commitments within normative ethics, but as soon as we start to spell out a particular relativist view, we get such commitments. The general form that a relativist view takes, is to tell us, for each context C and perspective P, whether some sentence S expresses a claim in C that is true with respect to P. But as soon as we have such a theory in hand, all that it takes to generate normative commitments is to know what context we are in and what our perspective is. If you know ‘my context is C’ you can get from ‘ ‘stealing is wrong’ expresses a claim in C that is true with respect to P just in case S’ to ‘ ‘stealing is wrong’ expresses a claim in my context that is true with respect to P just in case S’. But in general, as before, claims of the form ‘ ‘R’ expresses a claim in my context that is true with respect to P just in case it is true that R with respect to P’ express a truth in every context, and so we are committed to the claim, ‘ ‘stealing is wrong’ expresses a claim in my context that is true with respect to P just in case it is true that stealing is wrong with respect to P’. So we can infer, ‘it is



true that stealing is wrong with respect to P just in case S'. But 'it is true that stealing is wrong with respect to P just in case stealing is wrong' expresses a truth with respect to perspective P. So if we know that our perspective is P, we must accept this claim, and hence we may infer, 'stealing is wrong just in case S'. So again, we get a commitment to highly general normative consequences, merely by knowing our theory, knowing our context, and knowing our own perspective.

The reason why contextualist and relativist theories, once they amount to concrete theories and not just the bare assertion of contextualism or relativism, are committed to normative consequences, though traditional noncognitivist theories are not, is simple. Each class of theory purports to tell us something about the meaning of moral words. To do so, each theory mentions moral words or sentences involving moral words, and says things about those sentences. But the contextualist and relativist say things about moral sentences from which some moral sentences themselves follow, given appropriate, and available, background assumptions. That is because the things they tell us about moral sentences include things like 'stealing is wrong' is true iff P'. But for speakers who speak the language in which S is formulated, from this follows 'stealing is wrong iff P'. Noncognitivists do not use semantic vocabulary that is disquotational in this way. From 'stealing is wrong' is used to condemn stealing' no sentence that *uses* 'wrong' follows. This is because 'is used to' does not have any disquotational properties.

#### 4 Reductive Naturalism

Like contextualism and relativism, (noncontextualist) reductive naturalism is an important class of views in metaethics that cannot avoid substantive, deep, commitments within normative ethical theory. According to reductive naturalism, moral properties like *wrong* reduce to or are analyzable in terms of ordinary, non-moral, naturalistically respectable properties like desire, promotion, and explanation (Schroeder [2007]).

Of course, as with contextualism and relativism, it is possible to endorse the bare thesis of reductive naturalism without making any substantive normative ethical claims. It is even possible to give arguments that reductive naturalism must be true (as in Jackson [1998]) or to defend reductive naturalism from at least some important objections without taking on any such commitments. But as with contextualism and relativism, any complete statement of a reductive naturalist theory must tell us *how* moral properties reduce in naturalistic terms.

In fact, it is particularly pressing for the reductive naturalist to defend a particular reductive naturalist view, or at least a view about the neighborhood in which such a view will lie, and not simply to argue for reductive naturalism in the abstract. This is because reductive naturalism is often discussed by imagining

particular reductive views that are clearly false. Everyone, even advocates of reductive naturalism, should agree that at most one particular reduction of moral properties in naturalistic terms is true (compare Finlay [2014] on the Open Question Argument). So if a reductive naturalist refuses to take on the task of defending a *particular* reductive view, she leaves discussion of reductive naturalism in the hands of her critic, who may choose to illustrate it with the example of views that everyone can agree are not true.

Michael Huemer, for example, argues against reductive naturalism like this:

On the face of it, *wrongness* seems to be a completely different *kind* of property from, say, *weighing 5 pounds*. In brief:

1. Value properties are radically different from natural properties.
2. If two things are radically different, then one is not reducible to the other.
3. So value properties are not reducible to natural properties.

[...] To illustrate, suppose a philosopher proposes that the planet Neptune is Beethoven's Ninth Symphony. I think we can see that that is false, simply by virtue of our concept of Neptune and our concept of symphonies. Neptune is an entirely different kind of thing from Beethoven's Ninth Symphony. No further argument is needed. [Huemer 2005, 94]

Huemer argues against reductive naturalism by observing that an obviously false view is false, and then generalizing. But arguments do not need to be as bad as Huemer's in order to illustrate this point; indeed, the *best* objection to reductive naturalism is that no one has offered a proposed reduction that seems like it could possibly be true. High-level arguments that some reductive view *must* be true simply do not address the source of skepticism that no such view *could* be. The only antidote to this – the only dialectically fruitful way forward – is to defend better views, and to show that they are not obviously false.<sup>5</sup>

In fact, other objections to reductive naturalism actually turn on skepticism about whether there is any plausible naturalistic reductive analysis. The most famous of these is Terry Horgan and Mark Timmons' [1991] Moral Twin Earth argument. The Moral Twin Earth argument is based on a thought experiment that is designed to test and provide intuitive counterexamples to theories of reference-determination for moral words. But in its full generality, it claims to provide reason to think that *no matter what* reductive analysis

---

<sup>5</sup> Nick Laskowski [ms] argues, quite strikingly, that if reductive realism is true, then the correct reduction is not knowable. He contends that reductivists should restrict their ambition to claiming to know that *some* reductive thesis is true, without being able to know which one, and that this perspective solves independent problems for reductive realism, including addressing the objection that if reductive realism is true, then moral concepts are dispensable (for which, see Parfit [2011]). However, my points in this section can survive Laskowski's idea, provided that we can at least make progress in narrowing down the right *kind* of reductive theory.

we give to moral properties, and *no matter what* theory we have about in virtue of what moral words pick out those properties, there will be intuitive counterexamples to that combination of theories.

It is very puzzling, however, what could lead anyone to think that there will be intuitive counterexamples to *every* possible theory of reference-determination for moral words, without surveying every possible such theory. After all, in order for the counterexamples to be intuitive, there must be some features of the case that pump the intuition in question. But then a different theory of reference-determination could make those features relevant. In general, once we know what features of cases pump the intuition that speakers mean *wrong* by ‘wrong’, we can design our theory of reference-determination so that it assigns *wrong* as the meaning of ‘wrong’ under precisely those conditions. This makes it very puzzling indeed how anything general could follow from the structure of Horgan and Timmons’ argument.

But although they claim that their argument is an argument against any reductive theory, Horgan and Timmons help themselves to a helpful stipulation, in order to pump intuitions. They assume, for purposes of their argument, that the correct analysis of what we mean by ‘wrong’ makes it pick out what they call a ‘consequentialist property’. So they are stipulating for convenience that the reductive naturalist view under consideration claims something like that to be wrong is to fail to maximize happiness. Similarly, they stipulate that speakers on ‘moral twin earth’ use the word ‘wrong’ to pick out a ‘deontological property’. Since we know that consequentialists disagree with deontologists, this serves to pump the intuition that we disagree with speakers on moral twin earth, and hence that they also mean *wrong* by ‘wrong’. An alternative diagnosis, however, is that either the consequentialist analysis is wrong – perhaps because it is insufficiently charitable to the deontologists among us – or that moral twin earth would have to be much more different from ours than Horgan and Timmons go to pains to make us think, in order for speakers there to mean something else by ‘wrong’ – so different that it would not be intuitive, after all, that we disagree with them, once the differences are made vivid.

So I infer that one of the most pressing tasks facing the reductive naturalist is to defend actual reductive, naturalistic analyses of moral properties like *wrong*. Such an analysis will tell us *what it is* to be wrong. But once we know what it is to be wrong, we will know a condition that is necessary and sufficient for something to be wrong – indeed, we will know the most fundamental and explanatory such condition, in virtue of which all other generalizations about what is wrong are true. So we will have not just *some* commitment or other within normative theory, we will have attained the holy grail of the most ambitious and general kind of explanatory normative ethical theory. So the commitments of reductive naturalism for normative ethics run deep.

## 5 Ambition in Normative Theory

Indeed, I think the relationship between reductive naturalism and ambitiously general theorizing within normative ethics runs in both directions. This because general, explanatory normative ethical theories can take one of two forms, and of these two forms, the one with the better claim to generality and explanatory power is committed to a reductive thesis.

To see the distinction between the two forms of general, explanatory, normative ethical theory, compare two possible interpretations of consequentialism. According to the first interpretation, endorsed by Sidgwick [1907], there is only one act-type that is universally obligatory, and that is maximizing the good. No matter what your situation, you are morally obligated to maximize the good. And insofar as you are obligated to do anything else, that is because given your situation, that is what you must do in order to fulfill your obligation to maximize good. So on this view, facts about what will maximize the good explain what you are obligated to do because they are the features of your circumstances that affect what you need to do in order to fulfill your basic obligation. On this interpretation, consequentialism differs from deontological theories that consist in basic lists of obligations only by having fewer obligations on its list.

In contrast, Moore [1903] endorsed a strikingly different interpretation of his consequentialism. According to Moore's form of consequentialism, the reason why facts about what will maximize the good explain what you are obligated to do is that *what it is* for an action to be obligatory is for it to maximize the good. On this view, there is no basic obligation to maximize the good at all – only a fact about what *obligation* (Moore himself framed the view in terms of *rightness*) consists in.

Setting aside the act of maximizing the good, these two interpretations of consequentialism agree about what explains why obligatory actions are obligatory, though they disagree about how this explanation works. Each agrees that for every action that is obligatory for any agent in any circumstances except for maximizing the good, this action is obligatory for that agent in those circumstances because given those circumstances, unless the agent performs that action, she will not maximize the good. So over the scope of most of normative ethical theory, these two interpretations of consequentialism agree about a lot. But with respect to the act of maximizing the good, they disagree. Sidgwick's form of consequentialism offers no explanation at all of why the act-type of maximizing the good is obligatory. On the contrary, that it is obligatory is an unexplained posit of the Sidgwickian theory. So the Sidgwickian theory, though it offers a *highly* general explanatory theory, is not *perfectly* general. In contrast, Moore's form of consequentialism *is* perfectly general. According to the Moorean form of consequentialism, *if* it is obligatory to maximize the

good, that is because compared to all of its alternatives, the act of maximizing the good will result in better consequences than the alternatives.<sup>6</sup>

So of these two forms of consequentialism, the Moorean form offers an explanation that is strictly more general. That means that it is motivated purely from an ambition for generality in our normative ethical theorizing. But the way that the Moorean view works is by making a *constitutive* claim, about *what it is* for an action to be obligatory. And such claims are intuitively the province of metaethics – they tell us about the nature or analysis of some moral property.

Of course, Moore's analysis does not make him a reductive naturalist. On the contrary, Moore's analysis treats *right* (and derivatively, obligatory) in terms of *good*, itself a further normative property. But now the very same questions arise in ambitiously general explanatory normative theory of the good, as arose in such theory about the right. We can ask for general theories about what is good and why, just as we can ask for general theories about what is right and why. As it happens, whereas Moore advocated a powerful and general explanatory theory about right action, he was content in the case of the good, to settle for a list. With respect to the good, Moore takes the same sort of view as Clarke, Price, and Prichard take about the right.

There is nothing inconsistent about Moore's position. But if we were attracted to his view about right action because of its ambitious explanatory power, then we will be attracted to theories of the good with similar scope and generality. And as with consequentialism, these will come in two forms: those that advert to a list of one or more basic intrinsic goods, such that everything else that is good is good because it results in one of the things on this list, and those that tell us *what it is* for something to be good.

For example, consider the view that things are good because there are reasons to prefer them. On one interpretation, this view claims that what is intrinsically good is there being something that there are reasons to prefer. What makes happiness good, on this interpretation, is that given our circumstances, we have reasons to prefer it, and hence by getting happiness, we are getting something that is of genuine intrinsic value – namely, having something that we have reasons to prefer. On a different interpretation, however, this view claims that *what it is* for something to have intrinsic value is that there are reasons to prefer it (Korsgaard [1983]). Like Moore's form of consequentialism, this view does not claim that there being something that there are reasons to prefer is itself good (this is a substantive question), and it does not use this claim to

---

<sup>6</sup> If this sounds trivial to you, compare the same claim about the Divine Command Theory, according to which an act is obligatory just in case and because it was commanded by God. The Sidgwickian Divine Command Theory appeals to a basic obligation to obey God, while the Moorean version claims that the fact that God has commanded something is what it is for it to be obligatory (compare Schroeder [2005]). The Moorean may or may not think that it is obligatory to obey God; whether this is so will depend, on her view, on whether God has commanded us to obey Him. Similarly, even for consequentialism, whether the Moorean consequentialist thinks that it is obligatory to maximize the good will depend on details about what it means to maximize consequences.

explain why other things are good. Like Moore's form of consequentialism, it is strictly *more general* in its explanatory power. And like Moore's form of consequentialism, it adverts to a constitutive or reductive claim about *what it is* for something to be intrinsically good.

Again, this view (as it happens) is not reductively naturalistic. But if we continue down this path, then at some point we will run out of further moral or normative concepts to appeal to, in making our constitutive claims. At that point, in order to follow Moore's path over Sidgwick's, the constitutive account that we appeal to *will* have to be reductively naturalistic. But the Moorean path can be motivated over the Sidgwickian path purely on grounds of greater explanatory generality. And greater explanatory generality is an ambition that makes perfect sense even purely within normative ethical theory. Hence, pressure toward accepting a reductive naturalistic view can be motivated purely within normative ethical theory.

Shelly Kagan puts a similar thought this way:

[A]s we go deeper in our attempt to articulate the fundamental moral principles, relatively specific first-order claims about the content of morality give way gradually to more general overall characterizations of morality's content; and as these in turn become more general still, we find ourselves making what increasingly come to seem like second-order claims about the very nature of morality. This is especially so when we attempt to provide a foundation for the substantive moral claims of normative ethics. Such foundational theories will inevitably grow out of and appeal to larger metaethical conceptions of morality's purpose and point. That is, in the course of defending a given theory about the foundations of normative ethics, when we try to explain why it is that the various features of that theory should seem attractive and plausible, inevitably the claims we make will themselves simply be metaethical claims about the nature of morality. At a deep enough level, normative ethics does not merely draw upon metaethics – it simply becomes metaethics. (Kagan [1997, 5-6])

## 6 Summary

Though distinct in their primary questions and concerns, metaethics and normative ethics are not fully independent. Indeed, it follows from what I have argued that they are deeply intertwined enterprises, to such an extent that it is not surprising that it is sometimes hard to say of some philosophical issue, such as particularism, whether it belongs to normative ethics or to metaethics. Still, some interesting and central metaethical views can be stated in full without adopting any commitments in normative ethics. These theories were the ones being advocated most often in the middle of the twentieth century, by those who insisted most sharply on the distinction between metaethics and normative ethics, so it is no surprise where we came by the idea that these closely related domains of inquiry can be so easily separated.<sup>7</sup>

---

<sup>7</sup> Special thanks to Tristram McPherson, David Plunkett, Caleb Perl, Jonas Olson, Lea Schroeder, and especially to Nick Laskowski.

## References

- Boisvert, Daniel [2008]. 'Expressive-Assertivism.' *Pacific Philosophical Quarterly* 89(2): 169-203.
- Brogaard, Berit [2008]. 'Moral Contextualism and Moral Relativism.' *Philosophical Quarterly* 58(4): 385-409.
- Copp, David [2001]. 'Realist Expressivism: A Neglected Option for Moral Realism.' *Social Philosophy and Policy* 18: 1-43.
- Finlay, Stephen [2014]. *A Confusion of Tongues*. New York: Oxford University Press.
- Fisher, Andrew [2014]. *Metaethics: An Introduction*. New York: Routledge.
- Hay, Ryan [2011]. 'Hybrid Expressivism and the Analogy Between Pejoratives and Moral Language.' *European Journal of Philosophy* 21(3): 450-474.
- Horgan, Terence, and Mark Timmons [1991]. 'New Wave Moral Realism Meets Moral Twin Earth.' *Journal of Philosophical Research* 16: 447-465.
- Huemer, Michael [2005]. *Ethical Intuitionism*. New York: Palgrave MacMillan.
- Jackson, Frank [1998]. *From Metaphysics to Ethics*. Oxford: Oxford University Press.
- Kagan, Shelly [1997]. *Normative Ethics*. New York: Westview Press.
- Korsgaard, Christine [1983]. 'Two Distinctions in Goodness.' *The Philosophical Review* 92(2): 169-195.
- Laskowski, Nicholas [ms]. *On Moral Thought*. Doctoral Dissertation, University of Southern California.
- Loeb, Don [2007]. 'The Argument from Moral Experience.' *Ethical Theory and Practice* 10(5): 469-484.
- MacFarlane, John [2014]. *Assessment-Sensitivity: Relative Truth and Its Applications*. Oxford: Oxford University Press.
- Mackie, J.L. [1977]. *Ethics: Inventing Right and Wrong*. New York: Penguin.
- Moore, G.E. [1903]. *Principia Ethica*. Cambridge: Cambridge University Press.
- Olson, Jonas [2014]. *Moral Error Theory: History, Critique, Defense*. Oxford: Oxford University Press.
- Parfit, Derek [2011]. *On What Matters*, volumes 1 and 2. Oxford: Oxford University Press.
- Perl, Caleb [ms]. *Deontic Positivism*. Doctoral Dissertation, University of Southern California.
- Ridge, Michael [2014]. *Inpassioned Belief*. Oxford: Oxford University Press.
- Schroeder, Mark [2005]. 'Cudworth and Normative Explanations.' *Journal of Ethics and Social Philosophy*.  
[www.jesp.org](http://www.jesp.org).
- \_\_\_\_\_ [2007]. *Slaves of the Passions*. Oxford: Oxford University Press.
- \_\_\_\_\_ [2009]. 'Hybrid Expressivism: Virtues and Vices.' *Ethics* 119(2): 257-309.
- \_\_\_\_\_ [2013]. 'Tempered Expressivism.' *Oxford Studies in Metaethics* 8: 283-314.
- Sidgwick, Henry [1907]. *The Methods of Ethics*, 7<sup>th</sup> edition. Indianapolis: Hackett.
- Strawson, P.F. [1950]. 'On Referring.' *Mind* 59: 320-344.

Toppinen, Teemu [2013]. 'Belief in Expressivism.' *Oxford Studies in Metaethics*, volume 8, 252-282.