

## Semantic Naturalism and the New Naturalistic Fallacy

### I

More than a century ago, G. E. Moore famously offered an extended inference to reject what are in effect two substantially different types of ethical naturalism. Although some naturalistic doctrines targeted by that inference make semantic claims that, if true, would entail certain metaphysical claims, it is also possible that those semantic doctrines could be false and the metaphysical ones true at the same time. For if semantic naturalism is true, then moral terms and sentences are reducible, by an analysis of what they mean, into some purely descriptive terms and sentences. But, if metaphysical naturalism is true, then the moral properties and propositions expressed by moral terms and sentences are reducible to purely natural properties and propositions.

By conflating claims about meaning with claims about reference, Moore's original inference proved vulnerable to a number of well known objections. But we believe there is a revised Moorean inference that can succeed, and we offer that argument here. Its scope is restricted to *semantic reductivist* forms of ethical naturalism: those that rest on theses claiming the reducibility of expressions in the moral vocabulary to expressions in a non-moral one. A not uncommon philosophical view in Moore's day, such theses remain live options in metaethics (as we'll presently show). Although our inference is run against some putative equivalences between moral and purely descriptive terms, it could easily be adapted against ethical naturalist theses involving moral concepts and thoughts as well as sentences. Like Moore, we take the successful inference to consist in an open question argument ('*OQA*') followed by the naturalistic-fallacy

charge (*NFC*). But, unlike Moore, we identify *semantic naturalism* as the only naturalistic doctrine vulnerable to the inference, and we offer adequate support for the inference's premises. Naturalism in ethics is thus undermined, though not refuted in all its forms: semantical naturalism comes out false, but metaphysical naturalism might still be true.

## II

Among Moore's aims in his extended inference was to show that good is *sui generis*: that it is a unique and peculiar property not reducible to any naturalistic or metaphysical property or set of properties. Our concern here, however, is not principally with Moore's moral ontology. Rather, we take his extended argument to aim at refuting any attempt to reduce moral terms to purely descriptive terms. As many philosophers have noted, Moore's inference has in fact no tendency to undermine naturalistic attempts to reduce moral *properties*, though Moore himself plainly took these to be refuted by his anti-reductivist argument about moral terms. Thus, his *OQA* comes out stronger when (1) it is reconstructed as an argument against naturalistic semantic analysis of moral terms (often referred to by Moore as 'definition'), and (2) it takes 'good' as a paradigm moral predicate for intrinsic value. The argument then runs,

### *OQA*

- 1 Suppose that moral terms are analyzable into purely descriptive terms.
- 2 If moral terms are analyzable in to purely descriptive terms, then 'good' is analyzable into a purely descriptive predicate such as 'pleasure maximizing.'
- 3 If 'good' is analyzable into 'pleasure maximizing,' then the question, 'Granted *a* is pleasure maximizing, but is it good?' is closed.

4 But the question, ‘Granted *a* is pleasure maximizing, but is it good?’ is open.

*Therefore,*

5 ‘Good’ is not analyzable into ‘pleasure maximizing.’

6 Steps (1) through (5) can be iterated for any putative descriptive analysis of  
‘good.’

*Therefore,*

7 ‘Good’ is not analyzable into any purely descriptive terms.

Needless to say, if (7) is true, then ethical naturalism, as conceived by Moore, must be false.

Crucial to that conclusion, however, is premise (6), which generalizes over *any* attempt to analyze ‘good’ in descriptive terms. And what reasons, if any, does Moore offer for this premise? After all, unbeknownst to him there might be a *correct yet purely naturalistic* analysis of ‘good.’

To meet this objection, Moore could reply that the generalization in (6) is supported by cases, where these consist in a representative number of candidates for the most plausible naturalistic reductions. (In his time, these included predicates such as ‘better evolved’ and ‘what we desire to desire.’)<sup>1</sup> If for any proffered naturalistic definition of ‘good’ the question appears to remain open, then at length we abandon the attempt. Once premise (6) is supported, then, provided the argument leading to (5) is compelling (more on this later), (7) follows, and Moore is

---

<sup>1</sup> On our view, given the semantic nature of the issues involved in *OQA*, it must be reconstructed as having deductive form, even when some of its premises may be supported by abduction. For example, if premise (4) is supported by invoking speakers’ intuitions, then support for it is abductive. Cf. Ball 1991 (pp. 37-8), who considers it neither deductive nor inductive. Ball places it within “linguistic philosophy” tradition, and takes it to consist in a series of linguistic tests that are inductive inferences from linguistic behavior. See also Pigden 2007 and Snare 1975.

in a position to run the last part of his extended inference – namely, the naturalistic-fallacy charge:

*NFC*

8 If (7), then ‘good’ stands for a *sui generis* moral property.

From (7) and (8) it follows that

9 ‘Good’ stands for a *sui generis* moral property.

10 If (9), then any attempt to analyze ‘good’ into purely descriptive predicates commits the naturalistic fallacy.

*Therefore,*

11 Any attempt to analyze ‘good’ into purely descriptive predicates commits the naturalistic fallacy.

Although this extended inference focuses on ‘good,’ similar inferences are available for other predicates of intrinsic value or disvalue. Furthermore, similar inferences would be available for ‘right,’ ‘wrong,’ ‘ought to be done,’ and other deontic predicates that Moore took to be analyzable in terms of ‘good’. From this, further conclusions would follow – viz.,

12 No moral predicate is analyzable into purely descriptive predicates.

13 Any attempt to analyze a moral predicate into purely descriptive predicates commits the naturalistic fallacy.

### III

It is now notorious, however, that Moore’s *OQA* falls short of providing conclusive reasons for his own metaethical view, nonnaturalism, for it seems to *prove too much*: its

conclusion gives support, not only to nonnaturalists, but equally to the rival views of noncognitivists and error theorists. Those who hold either that moral judgments are *not descriptive*, or else that they *are* descriptive but *have no referents*, can both accept Moore's *OQA* and the related notion of a naturalistic fallacy. Yet the incompatibility of these doctrines with Moore's nonnaturalism becomes apparent when we consider the epistemic attitudes they imply toward *sui generis* nonnatural moral properties: the nonnaturalist professes belief in their existence, the error theorist professes disbelief (i.e., believes that *it is false* that there is any such property – in effect, embracing nihilism), while the noncognitivist is at best agnostic, suspending judgment on whether there are such properties. Thus Moore's original argument fails, on the one hand, to support a strictly nonnaturalist moral ontology, but also, on the other, falls short of undercutting all versions of ethical naturalism.

If we consider briefly the theses available to reductive programs of ethical naturalists it becomes evident why this is so. A motivation for naturalism in ethics is of course *philosophical naturalism*, a worldview according to which *all there is* is the natural world, and all causes are natural causes. Given philosophical naturalism, there are neither supernatural entities nor supernatural causes. *Ethical naturalism*, on the other hand, is the view that all moral facts are natural facts. Plainly, this is consistent with philosophical naturalism but not entailed by it. By contrast, *ethical nonnaturalism* may or may not be consistent with philosophical naturalism, depending on how the former metaethical doctrine is construed. As proposed by the leading nonnaturalists of the early twentieth century, and also by current representatives, it may countenance an extravagant ontology but need not amount to supernaturalism.<sup>2</sup> On our view, it is

---

<sup>2</sup> Nonnaturalism would be in the clear if a case can be made out, as some current nonnaturalists seem to think, that moral properties are similar to abstract entities such as meanings and the objects of math and logic. See Shafer-Landau 2006.

not enough to construe nonnaturalism as simply *the rejection of ethical naturalism*, since the former is merely one among several metaethical views that do that. In addition to nonnaturalists, other ‘ethical anti-naturalists’ include supernaturalists, noncognitivists, and nihilists of various persuasions, such as error theorists and moral fictionalists. Each of these groups holds a quite distinct position different from the others – and different, too, from ethical nonnaturalism.

Yet it is not to any of these, but to the various forms of ethical naturalism that we must look if we wish to understand nonnaturalism. Some of these naturalist theses are reductive, others not. But, since the only doctrines vulnerable to a Moorean inference are reductive ones, let’s start with them. Standardly construed, reductive ethical naturalism can consist in a semantical, metaphysical, and/or epistemic doctrine, which we define as follows:

1. Semantical Naturalism (*SN*)

*SN1*

Moral terms and sentences are analyzable, without evaluative or normative remainder, into purely descriptive terms and sentences.

*SN2*

Moral concepts and thoughts are analyzable, without evaluative or normative remainder, into purely descriptive concepts and thoughts.

2. Metaphysical Naturalism (*MN*)

Moral properties are analyzable, without evaluative or normative remainder into purely natural properties.

3. Epistemic Naturalism (*EN*)

Inquiry into the nature of morality needs no justification beyond that attainable through the standard methods of science.

Doctrines such as metaphysical and semantical naturalism seem to have the consequence of epistemic naturalism, which by itself it does not entail the repudiation of ethics, but rather its assimilation into the sciences. That is, epistemic naturalism challenges the autonomy of ethics as a discipline with its own subject matter and methods. To nonnaturalists, by contrast, the autonomy of ethics follows from the *sui generis* nature of moral predicates, something Moore sought to prove with the *OQA*.<sup>3</sup>

As suggested above, however, the *OQA*, under the most charitable reconstruction, is at most an argument against semantic versions of ethical naturalism. These could be made to consist in either *SN1* (a thesis about the linguistic meaning of moral terms and sentences) or *SN2* (a thesis about moral concepts and judgments). Either way, theses along these lines are the most vulnerable to Moore's inference. Moore himself seems to have had in mind *both* theses, given that he sometimes objects to naturalistic definitions of the meaning of moral predicates, while his concern on other occasions is with naturalistic conflations of the moral and purely descriptive concepts as they appear 'before the mind.' In any case, there is an independent reason for taking

---

<sup>3</sup>Moore's arguments against ethical naturalism are offered in the course of his reflections on the subject matter of ethics (esp. in Chapter 1 of *Principia Ethica* and in the *Lectures*), where he shows concern with vindicating the autonomy of ethics (principally but not exclusively) from psychology and theory of evolution. At the same, Moore rejected any variety of supernaturalism and thus upheld the autonomy of ethics from religion. See also C. D. Broad 1949.

*SN1* and *SN2* to be parallel: the plausible assumptions that when speakers are sincere and competent, what they mean is what they think, and the so-called Vendler's Hypothesis (viz., that reports of what a person means and what she thinks have a parallel syntactic structure).

In light of this parallel, we'll hereafter construe semantical naturalism simply as *SN1*. Would the falsity of *SN1* entail the falsity of metaphysical naturalism, as Moore seems to have thought? Clearly not, for although there is no consensus about the correct semantics for moral terms, on anyone's view, semantical naturalism could be false and metaphysical naturalism true. Moore's mistaken diagnosis here<sup>4</sup> seems to be the result of his conflating claims about the meaning of moral predicates with claims about the properties they denote. Certainly, no such conflation would have occurred if *all there were* to the meaning of certain moral terms and some putatively equivalent purely descriptive terms were their *denotation*. But no one today would accept such an implausible 'Fido'-Fido semantic account of those terms. Critics of various persuasions in philosophy of language are now agreed that, since predicates instantiating different semantic types may nonetheless denote the *same properties*, therefore, for all *OQA* shows, metaphysical naturalism might be true even if semantical naturalism is false. Thus Moore's actual inference (which, unlike the one above, equivocates between predicates and properties) fails precisely because it attempts to show the falsity of metaphysical naturalism from the failure of semantical premises alone.

But this naturalist response to Moore, however successful, still falls short of actually supporting metaphysical naturalism. To see this, let's consider the three versions mentioned

---

<sup>4</sup> The objection was raised independently by Bob Durrant (1970), Hilary Putnam (1981) and Gilbert Harman (1977). We have argued elsewhere (2007) that there are other reasons, besides a successful *OQA* against semantic naturalism, that jointly point to a substantial difference in the kinds of properties at issue in metaphysical naturalism.

earlier. All three naturalistic theses above are reductive: *SN* because, if true, then moral terms and purely descriptive terms are synonymous, which amounts to the claim that what could be meant by invoking the former could *equally* be meant by invoking the latter; *MN* because, if true, then moral properties just *are* nothing over and above purely natural properties; and *EN* because, if true, then ethics lacks a subject matter of its own, becoming one or more chapters of another discipline, such as psychology, sociology, or evolutionary science.

However, while semantic naturalism seems essentially reductive, neither metaphysical nor epistemic naturalism is limited in this way to a single option. Non-reductive metaphysical and epistemic theses are available to ethical naturalists – for example, one holding that moral properties supervene on purely natural properties. A naturalist who subscribes to this view would say that it's in virtue of having certain natural properties that an action has certain moral properties such as *rightness*, *wrongness*, and so on. The fact that Moore himself clearly held a supervenience view similar to this might make us suspect at first that he was actually a kind of non-reductive naturalist after all. But this would be to confuse two very different sorts of supervenience doctrines, for non-reductive naturalists hold that the moral properties that supervene on natural properties *are themselves natural properties*, while Moore's supervening goodness is of course a nonnatural one. Even so, the details of Moore's view on this question are often exceedingly difficult to nail down, as can be seen from his discussion in "The Conception of Intrinsic Value," where he struggles to draw a proper contrast between 'values' and natural predicates and properties (another distinction that we need to make at the outset). Close to the letter of Moore's argument, we say that intrinsic value supervenes on *natural intrinsic properties* of the action or thing that has the value. By the nonnaturalists' criterion, these properties are

properties that are necessary for any complete description of an action or object. That is, a Moorean natural property might be defined as follows:

Moorean Natural Property (*MNP*)

An intrinsic, basic property of an action or object is natural if and only if no description of that action or object could be complete if it omitted that property.<sup>5</sup>

Let's suppose that thin ethical concepts denote basic moral properties (leaving out Moore's requirement that they be "intrinsic, but not intrinsic *properties*"). By the above criterion, basic moral properties can contribute nothing to the description of the action or object that instantiates them. But neither can supernatural properties (if there are any), so the criterion is really too broad. For our purposes, let's define natural property this way:

Natural Property (*NP*)

A property counts as *purely natural* if and only if any claim about whether something has it is defeasible, at least in principle, by empirical evidence.

This criterion seems to capture what is essential to natural properties, partly because it relies on a clear idea of what it is that philosophical naturalists generally reject about doctrines that make claims not open to empirical tests – whether they involve meanings, demons, or nonnatural moral properties. On the other hand, it also accommodates the nonnaturalists' rejection of empirical methods in ethics: moral properties are somehow beyond those that are defeasible by empirical

---

<sup>5</sup>For Moore, only natural properties, he says, are eligible to be counted as intrinsic properties, given that they "seem to describe the intrinsic nature of what possesses them in a sense in which predicates of value never do. If you could enumerate all the intrinsic properties a given thing possessed, you would have given a complete description of it, and would not need to mention any predicates of value it possessed; whereas no description of a given thing could be complete which omitted any intrinsic property"(1922: 273). Moore provided nothing more precise that could allow us to draw a distinction between moral and natural properties, a problem that he continued to wrestle with long after *Principia* (see, e.g., his replies to some critics in Schilpp 1942).

test. Furthermore, if moral properties were purely descriptive properties, then, like any other natural properties, they would be amenable to study with the methods of the empirical sciences.

For nonnaturalists, to call a thing ‘good’ is obviously not to ascribe to it a natural property, but rather, a unique and peculiar property – even though the thing may have it in virtue of also having certain other, purely natural properties. There is of course considerable uneasiness among philosophers with the notion of a ‘nonnatural property,’ and nonnaturalists are sometimes accused of metaphysical and epistemic extravagance, or worse.<sup>6</sup> There is logical space to object that, if there were mind-and-language-independent *nonnatural* moral properties, these would be ‘queer,’ Platonistic, or even supernatural. Although such objections have been played down with the current revival of nonnaturalism, more attention needs to be paid to their assessment.<sup>7</sup>

In any case, we are now in a position to provide a metaphysical thesis that can distinguish nonnaturalism from other anti-naturalist doctrines. It holds that,

Metaphysical Nonnaturalism (*MNN*)

- (1) There are mind-and-language-independent moral properties and facts that are  
irreducible to natural properties and facts, and
- (2) each of these is *sui generis*.

*MNN* captures two claims of nonnaturalists: (1) the irreducibility claim (that moral properties and facts cannot be reduced to purely natural properties and facts), and (2) the moral

---

<sup>6</sup> For a recent discussion of this objection, see P. Stratton-Lake 2002.

<sup>7</sup> Deflationary accounts of this objection can be found, e.g., in Shaver 2007 and Hurka.2003.

realism claim.<sup>8</sup> But nonnaturalism may also consist in semantical and epistemic theses, each of which would be incompatible with *SN* and *EN* above:

Semantic Nonnaturalism (*SNN*)

Moral terms are not analyzable without evaluative or normative remainder into purely descriptive terms.

Epistemic Nonnaturalism (*ENN*)

The standard methods of empirical science are insufficient for inquiry into the nature of moral properties.

#### IV

In support of all three theses of nonnaturalism, Moore offered the *OQA*, followed by the *NFC*, which we construed here as an extended inference. Of these two parts of the inference, it is only revised versions of the *OQA* that are sometimes vindicated in the literature: the naturalistic-fallacy doctrine, by contrast, usually finds no takers at all. Yet we believe that, properly construed, both parts of the inference are in fact defensible. Needless to say, such a defense must provide an independent argument to the effect that the sort of mistake committed by naturalistic analyses amounts to a *fallacy*, since a compelling *OQA* is necessary but not sufficient to substantiating the *NFC*. Let's now consider in turn each of these parts of the inference.

First, as noted above, although Moore set himself to disprove *all* ethical naturalist doctrines with the *OQA*, it is only *some* versions of reductive naturalism that are vulnerable to his

---

<sup>8</sup>Although compatible with Jonathan Dancy's definition of nonnaturalism, we think that ours avoids the problem of being too broad. On Dancy's view, nonnaturalism is "the claim that ethical properties, distinctions, and facts are different from any properties, distinctions, and facts that are worth calling natural." (2006: 122).

argument: namely, those that attempt semantic analyses of normative terms into purely descriptive terms. If such reductive analyses were plausible, they would of course advance the agenda of those who seek to accommodate ethical value within a commonly held naturalistic worldview. The equivalence sanctioned by any such analysis would be strictly semantical – that is, it would sanction *the sameness of meaning* between tokens of normative and purely descriptive terms: any such tokens would instantiate the same semantic type. That's precisely what is sanctioned, for example, by the following reductive analysis:

*Hedonist*

‘Good’ = ‘pleasure-maximizing’

Although the revised *OQA* that we will propose here appeals to *Hedonist* as a paradigm case of naturalistic analysis, our inference is in no way dependent on the simplicity of this model. In our view, it is not only *Hedonist* but *any* semantic analysis of ethical naturalism that will be open to doubt on a priori grounds. Thus there is room to argue that analyses of that sort do commit a *fallacy* that may be labeled ‘naturalistic.’ Our revised version of Moore’s argument is an extended inference that runs as follows:

*OQA*\*

- 1 Suppose that moral terms are content-equivalent to purely descriptive terms.
- 2 If moral terms are content-equivalent to purely descriptive terms, then ‘good’ is content-equivalent to terms such as ‘pleasure maximizing.’
- 3 If ‘good’ is analyzable into ‘pleasure maximizing,’ then the question, ‘Granted *a* is pleasure maximizing, but is it good?’ is not debatable on a priori grounds.

4 But the question, ‘Granted *a* is pleasure maximizing, but is it good?’ is debatable on a priori grounds.

*Therefore,*

5 ‘Good’ and ‘pleasure-maximizing’ do not have the same content.

6 Steps (1) through (5) can be iterated for any attempted naturalistic analyses of the term ‘good’ into purely descriptive terms.

*Therefore,*

7 ‘Good’ is not content-equivalent to any purely descriptive terms.

8 Steps (1) through (7) can be iterated for any attempted naturalistic analyses of other normative terms into purely descriptive terms.

*Therefore,*

9 Normative terms are not content-equivalent to any purely descriptive terms.

*NFC\**

10 If (9), then normative terms are semantically *sui generis*.

*Therefore,*

11 Normative terms are semantically *sui generis*.

10 If (11), then any attempt to analyze normative terms into purely descriptive terms commits the naturalistic fallacy.

*Therefore,*

12 Any attempt to analyze normative terms into purely descriptive terms commits the naturalistic fallacy.

Given the parallel between linguistic meaning and mental content, a similar inference could be run to the effect that normative concepts are not semantically equivalent to non-normative concepts. Crucial premises in the *OQA\** component of this inference are (2) and (3), where Moore's own claims about the question's being closed *versus* its being open have been cashed out in terms of a priori knowledge. We have argued elsewhere for a quasi-Cartesian view of self-knowledge, according to which at least self-ascriptive, comparative judgments about mental content are a priori justified – in the sense that, under normal circumstances, their justification does not require investigation of the environment. As we see it, this view would be challengeable only if a case could be made that normative terms are similar to natural-kind words or proper names, in that their content involves their extension. But externalism about normative terms seems implausible (see @).

Now, to support the *NFC\** component of our extended inference, we need only add a straightforward consideration about dialectical space and some clear notion of the sort of mistake in reasoning that can be rightly counted as a *fallacy*. Consider *Hedonist*: to qualify as a fallacy, this analysis cannot rest simply on a *false claim of identity* between two predicates that are not so equivalent – even though the term, 'fallacy,' is sometimes used loosely to mean 'mistaken belief', as when someone speaks of the "fallacy that animals do not feel pain" (which seems to express the belief that a certain propositional content is defective, perhaps by being false, or at least unsupported by the evidence). If *that* is all that goes wrong when a semantic analysis of this sort is advanced, no *fallacy* would be involved in it. If, on the other hand, the mistake made in advancing semantic analyses of this kind is one that merits labeling and classification among the fallacies (in this case, clearly the so-called informal fallacies), then we need some clarity on just

what sort of classification this is. What, exactly, is meant by ‘fallacy’ when used as a term of art?

We submit that it can be defined roughly in this way:

#### Fallacy

A type of mistake in reasoning is a *fallacy* if and only if it amounts to a pattern of failed relation among concepts or beliefs affecting any reasoning that instantiates that pattern.

The relevant relation at issue in a fallacy may be conceptual, inferential, or explanatory. Some fallacies hinge on logical form, others on semantic and pragmatic factors. Of interest here are the latter, which, in the case of argument, amount to patterns of mistake in reasoning involving something other than invalidity. But the fallacy at work in *Hedonist* concerns definition and arises only in the context of plausible versions of the open question argument. In the context of such arguments, the semantic analyses of some ethical naturalists amount to proposed definitions that are dialectically abusive, and so the fallacy committed by them belongs to the pragmatic or dialectical informal fallacies, along with other familiar fallacies such as *petitio principii*, straw man, and complex question.

To see the pattern of mistake we have in mind here, suppose that, in the context of the debate over the moral status of abortion, you define ‘abortion’ as ‘a form of infanticide.’ This is then not merely an *opinion* of yours that others could regard as mistaken or in need of support. For, in the context of the abortion debate, your definition is clearly *abusive in some sense*; after all, it presumes to *settle by the meaning of its terms alone* what is in fact a disputed claim. Just as ‘killing fetuses’ and ‘killing babies’ are not obviously semantically equivalent, so the term ‘abortion’ does not obviously *mean the same* as ‘a form of infanticide.’ A similar mistake would

be to define ‘marriage’ as ‘the union between one man and one woman,’ ‘euthanasia’ as ‘medical murder,’ or ‘communism’ as ‘unjust system of government.’ Examples of this sort all share a certain *pattern* of defective reasoning, one that has sometimes textbooks refer to as the ‘definist’ fallacy:

### Definist Fallacy

The type of mistake made in attempting analytic definitions of terms whose semantic identity of content is, in a certain dialectical context, debatable.

Tokens of definition that are abusive in this way instantiate a type of fallacy because, in the context of a deliberation, they would fail to discharge the burden of proof. They are, in effect, begging the question *against* the opposing view. Confronted with this sort of tactic in an actual debate, a critical thinker can dismiss it rationally by simply pointing out that the definist fallacy has been committed.

Semantic analyses fueled by ethical naturalism (such as *Hedonist*) can now be shown to commit a species of definist fallacy that can be properly labeled the ‘naturalistic’ fallacy.<sup>9</sup> For let’s reconsider *Hedonist*, construed as a semantic analysis sanctioning sameness of meaning between a normative term, ‘good,’ and a non-normative term, ‘pleasure-maximizing.’ By proposing such an analysis, *Hedonist* commits a type of definist fallacy, given that the audience might come to have a priori reasons to *doubt* the proposed analysis. After all, a critical thinker could run *OQA*\*, thereby acquiring such reasons. That is, against the background of *OQA*\*, proposing *Hedonist* exemplifies the mistake of attempting to make true by definition a semantic

---

<sup>9</sup> Moore himself notes that the question of what *label* to adopt for this fallacy is unimportant. He writes, “And I do not care about the name: what I do care about is the fallacy. It does not matter what we call it, provided we recognize it when we meet with it. It is to be met with in almost every book in Ethics; and yet it is not recognized: and that is why it is necessary to multiply illustrations of it, and convenient to give it a name.” *Principia Ethica* §12, (1993: 65-66).

equivalence that is in fact a priori debatable. It is not difficult to see how other *OQA*\* arguments could be deployed for any attempted conceptual analysis of a value concept in terms of a purely descriptive concept – thus creating a context wherein each proposed analysis faces a similar charge. If we are right, then it is actually *semantic naturalists*, rather than their opponents, who in fact beg the question, for any of their reductive analyses seems vulnerable to a properly construed *NFC*\* grounded in an adequate *OQA*\*.

But if no questions are begged by the naturalistic-fallacy doctrine, why has it been thought to amount to a *petitio*? For one thing, critics have notoriously evaluated it in isolation from its dialectical context, and as result have failed to recognize the naturalistic fallacy as a species of a broader, dialectical phenomenon, the definist fallacy.<sup>10</sup> Their focus has been instead on metaphysical doctrines of reductive naturalism, which are *resistant* to arguments such as *OQA*\*. On the other hand, it is now clear that a defense of the naturalistic fallacy doctrine in the way we propose (that is, through the *OQA*\*) is in no way dependent on Moore’s inflationary view about the ethical analyses that have supposedly committed that fallacy, since naturalistic analyses of *properties* fall beyond the scope of our revised version of the argument.

At the same time, our inference can accommodate other, often criticized elements of Moore’s account, such as the notion that a type of mistake in reasoning similar to the naturalistic fallacy is made by some analyses of concepts that are both purely natural, or where one is ethical and the other metaphysical. On Moore’s own examples, to hold that “‘Goodness’ is *what is commanded by God*” commits the definist fallacy in the context of a suitable *OQA* that plausibly raises a priori doubt about whether those concepts are indeed equivalent. And although the

---

<sup>10</sup> Castings of the naturalistic fallacy that have gone wrong in these ways include, not only Frankena 1939, but also Williams 1985, Ridge 2006 and Darwall, Gibbard, and Railton 1997.

concepts on both sides of the analysis, “‘Pleasant’ is *the sensation of red*”, are purely natural<sup>11</sup>, it is not difficult to invoke an adapted *OQA*\* to generate a context that raises a priori doubt about their semantic identity – thus revealing a definist fallacy in the attempted analysis. But if the naturalistic fallacy is a species of the definist fallacy, then Moore was right in noticing that there is an obvious parallel between the type of mistake in the semantic analyses of ethical naturalists and the definist fallacy committed by similar attempted semantic analyses of other unsettled identities.

## V

One advantage of our construal of the naturalistic fallacy charge, we think, is that it makes evident just *what* the type of mistake might have been that was obscurely pointed out by Moore. It is now plain that that mistake must not be conflated with Hume’s famous observation that any argument attempting to derive normative conclusions from exclusively factual premises would be *invalid*. Standardly understood, the ‘No-Ought-from-Is’ rule sanctions the invalidity of any such attempt.<sup>12</sup> An argument often attributed to Mill<sup>13</sup> may be invoked here to illustrate how the Humean and the Moorean charges work. It runs,

### **EPICUREAN I**

P      Pleasure is sought by all people.

---

<sup>11</sup> As noted by Moore in *Principia Ethica* chapter 1.

<sup>12</sup> William Frankena (1939: 57) was among the first to distinguish between the ‘No-Ought-from-Is’ rule and Moore’s own naturalistic fallacy charge. But the distinction can be also found in White 1958 and Sylvester 1990. See Williams 1985 and Ridge 2006 for a different take on the issue.

<sup>13</sup> Whether Mill did actually argue this way has, of course, been a matter of considerable dispute.

*Therefore,*

C     Pleasure is good.

Although it might appear that **EPICUREAN I** flouts the No-Ought-from-Is rule, as Frankena points out, the argument does not fail on this count, since it is clearly enthymematic: it has a tacit premise that is itself normative. Once that premise is made explicit, the argument then runs as follows:

**EPICUREAN II**

P0     Whatever is sought by all people is good.

P1     Pleasure is sought by all people.

*Therefore,*

C     Pleasure is good.

Provided of course that there is no equivocation affecting the occurrences of ‘good,’

**EPICUREAN II** is obviously valid, and thus escapes any charge based on Hume’s rule. But it could still face the *NFC\**, depending on how its premise 0 is construed: when that premise is understood as stating an *analytic* identity, **EPICUREAN II** does face that charge. On the other hand, if premise 0 is read as stating a *synthetic* identity, then if there is empirical evidence supporting it, the argument would be sound (provided of course that P1 is also empirically supported). If there is no such evidence, then the argument would be noncogent, since premise 0 would beg the question against deontologists, ethical egoists and others who *deny* that whatever is sought by all people is good.<sup>14</sup> And it won’t do to argue that the identity in this premise *may be* parallel to the theoretical identities of science (‘Water is H<sub>2</sub>O,’ ‘Heat equals molecular motion,’

---

<sup>14</sup> For the accounts of begging the question that we have in mind here, see e.g. Jackson 1987, Davies 2003 and Wright 2007.

and so on). For plainly what would be needed is an argument to the effect that the identity *is* analogous to those synthetic identities.

But rather than a synthetic identity, premise 0 could be taken to express an analytic identity, true by the meaning of its predicates alone. Now the argument reads,

**EPICUREAN III**

P0' Whatever is sought by all people is good. (By definition)

P1 Pleasure is sought by all people.

*Therefore,*

C Pleasure is good.

**EPICUREAN III** is valid and may even be sound. Yet its premise 0' is vulnerable to the naturalistic-fallacy charge. As a result, the argument is dialectically ineffective, in the sense that it would fail to rationally persuade skeptical audiences to accept its conclusion, even after they've worked out the entailment. After all, those audiences could invoke an *OQA*\* to gain a priori grounds for doubting that premise.

**EPICUREAN III** would then be shown to fail, not because it flouts the Humean rule (the argument commits no logical error) but rather because it instantiates a type of mistake in definition that can occur only in a context of rational deliberation. The mistake in premise 0' consists in proposing a semantic equivalence that would be debatable on a priori grounds for a sufficiently critical audience. As it stands, **EPICUREAN III** would then fail to move that audience to accept its conclusion, even in cases where there are independent reasons to accept it.

Yet not all definitional premises would be vulnerable to the *NFC*\*. Consider the following (nonenthymematic) argument, which features one such premise:

## **HEALTH**

P1     Whatever is sought by all people is universally desired. (By definition.)

P2     Health is sought by all people.

*Therefore,*

C     Health is universally desired.

It may be argued that **HEALTH** fails, but not because of its vulnerability to the naturalistic-fallacy charge. To see this, let's compare it with **EPICUREAN III**: although both argument are valid, both fail to be cogent. But **HEALTH** is noncogent as a result of begging the question, **EPICUREAN III** as a result of committing the naturalistic fallacy. (On our view, to beg the question and to commit the naturalistic fallacy are both among the dialectical mistakes that fatally undermine cogency.) **HEALTH** begs the question, given that its premise 2 is not more acceptable than the conclusion it is offered to support. At the dialectical level, the argument then fails to support its conclusion for any audience who reasonably think that it is *not health*, but, say, *pleasure* that all people seek. In that dialectical context, proposing **HEALTH** begs the question.

By contrast, by competently and sincerely asserting **EPICUREAN III**'s premises, the arguer implicates that she has solid grounds for them, thus inviting the audience to accept them, work out the entailment, and secure justification for the conclusion. Since the grounds for premise 0' are a priori, an audience who run a suitable *OQA*\* would find the alleged semantic identity in that premise debatable also on a priori grounds. In other words, given the background beliefs of an audience sympathetic to the reasoning at work in *OQA*\*, premise 0' would be open to doubt. And since similar reasoning could be rehearsed for other semantic analyses of ethical naturalists, their being open to doubt on a priori grounds would seem not at all an accidental feature of a

particular audience and naturalistic definition of moral predicates in terms of non-moral ones. Rather, such definitions appear to exemplify a regular pattern of failure in the reasoning at work. We submit that this is what Moore should have responded to the objection that, even if the semantical analyses of some ethical naturalist are mistaken, they do not amount in any sense to a fallacy.

## VI

Finally, let's anticipate some further objections to our Moorean inference. We'll consider two main objections to Moore's *OQA*, and show how the revised version proposed here can meet them. These are the 'paradox of analysis' and straw man objections, which we shall take in turn. We have argued that answers to the Moorean question facing naturalistic analyses of moral terms trade on a priori thinking: to decide whether each alleged equivalence is an open or a closed question involves a first-person judgment about the contents of one's own moral terms. The question is open if and only if the alleged equivalence is *doubtable on a priori grounds*, and otherwise is closed. Cashed out this way, the status of the question does not depend on its being significant or non-trivial – as opposed to insignificant or trivial. Thus the inference featuring the Moorean question faces no paradox of analysis, given that on this reading a priori questions need not be trivial. Yet this reading still allows for a potentially challenging objection: after all, couldn't some naturalistic analyses be correct, but not trivial? If they are not trivial, they would be open to a priori doubts, and therefore, be sanctioned as 'open' by the Moorean test – which challenges the conclusions drawn by *OQA*\*. But the Moorean may concede that semantic naturalism cannot be proved false, while yet holding that *OQA* \* *makes its falsity plausible*.

Dialectical space here hinges on the status of intuitions about the meaning of one's own terms: even when such intuitions are neither self-evident nor incorrigible (that is, when they are defeasible by empirical evidence), under normal circumstances and in the absence of contrary evidence, they are direct and authoritative. Given their special epistemic status, the Moorean intuitions about specific semantic analyses of ethical naturalists appear *prima facie* more acceptable than those analyses – at least until semantic naturalists can make a compelling case for their correctness. To succeed against the Moorean inference, semantic naturalists must produce instances of such non-trivial analyses, together with arguments capable of overcoming the a priori doubts generated by that inference. More than claiming that such analyses are 'possible,' naturalists would need to *succeed in explaining away* the a priori doubts about them generated by suitable *OQA*\*s.

Now, what about the straw man objection? Our version of the Moorean inference targets only *semantic* versions of reductive naturalism in ethics – and, as such, was never vulnerable to the charge that the argument is a straw man, or that it amounts to flogging a dead horse. Abundant textual evidence supports this reply, for it shows that such semantic analyses are not uncommon in metaethics. Though Moore's claim to have found them in Mill is controversial,<sup>15</sup> he does seem to have found them in other ethical naturalists, such as Bentham, Spencer, and Westermarck.<sup>16</sup> Bentham did after all write that words such as 'right,' 'wrong,' and 'ought' *have meaning* only when interpreted in terms of the principle of utility, and also defines 'right action' as 'an action that is conformable to the principle of utility' (1789, ch. 1, §10 ff). Spencer took 'conduct that is

---

<sup>15</sup> See, e.g., Hall 1950: 51; Warnock 1960: 28-40; West 1997; and Darwall 1998: 114-15.

<sup>16</sup> Thomas Baldwin (1990: 69) notes that, if we go by what Moore says in *Principia*, he seems to have found the naturalistic fallacy in the work of Aristotle (§106), the Stoics (§67), Spinoza (§67), Rousseau (§27), Bentham (§14), Kant (§75), Mill (§40), Spencer (§32), and Green (§84).

more evolved' to be semantically equivalent to 'conduct that has a higher ethical sanction' (*PE* §29 and pp. 31-35); and Westermarck advocated a form of psychological naturalism according to which 'right' means 'arouses feelings of approbation' and 'wrong' means 'arouses feelings of disapprobation' (Moore 1922: 332). For unambiguous representatives of naturalistic analyses, however, Moore needn't have looked farther than the work of his own contemporaries, Ralph Barton Perry and F. C. Sharp. Perry (1926: 138) famously argued that "x is valuable = interest is taken in x," and Sharp (1928: 409-11) defined 'good' as 'desired upon reflection' and 'right' as 'desired when looked at from an impersonal point of view.'

In Moore's day, then, it appears that there was no shortage of analyses of the sort targeted by the naturalistic-fallacy doctrine. Even so, someone might reasonably ask, are these dinosaurs truly representative of metaethics *now*? Perhaps the semantic analyses of ethical naturalists have by now been discredited? Not at all – as can be seen from some current reductive programs attempting analyses of precisely that sort, such as Frank Jackson's (1998, 2003) and Michael Smith's (2000). For Jackson and Smith, there are cases in which, once a situation has been well described in *nonmoral terms*, certain *ethical sentences* will follow a priori from that description. As Jackson puts it, such ethical sentences are "a priori equivalent to and analyzable in terms of nonmoral ones" (2003: 558). In light of these current programs, no dead horses are being flogged by a properly construed Moorean inference.

## References

- Baldwin, T. *G. E. Moore* (London: Routledge, 1990).
- Ball, S. W. "Linguistic Intuitions and Varieties of Naturalism." *Philosophy and Phenomenological Research* 51, 1 (March 1991):1-38.
- Bentham, J. *An Introduction to the Principles of Morals and Legislation* [1789] (Buffalo, NY: Prometheus Books, 1988).
- Dancy, J. "Nonnaturalism," pp. 122-45 in *The Oxford Handbook of Ethical Theory*, ed. by D. Copp (Oxford: Oxford University Press, 2006).
- Darwall, S. *Philosophical Ethics* (Boulder, CO: Westview Press, 1998).
- Darwall, S. Gibbard, A., and Railton, P., eds. *Moral Discourse and Practice: Some Philosophical Approaches* (New York: Oxford University Press, 1997).
- Davies, M. "Begging the Question and Settling the Question." Hempel Lectures II, Princeton University, Sept. 2003.
- Dowden, B. "Fallacies," in *The Internet Encyclopedia of Philosophy* (2008).
- Durrant, R. G. 'Identity of Properties and the Definition of "Good,' *Australasian Journal of Philosophy* 48, 3 (December 1970): 360-61.
- Frankena, W. K. "The Naturalistic Fallacy," *Mind* 48 (1939). Reprint pp.54-62 in Sellars and Hospers.
- Hall, E. "'The 'Proof' of Utility in Bentham and Mill,' *Ethics* 61 (1950-51): 66-68.
- Harman, G. *The Nature of Morality* (New York: Oxford University Press, 1977): 17-20.
- Hurka, T. 'Moore in the Middle,' *Ethics* 113, 3 (Apr. 2003): 599-628.
- Jackson, F. *Conditionals* (Oxford: Blackwell, 1987).

----- . *From Metaphysics to Ethics: A Defence of Conceptual Analysis* (Oxford: Clarendon Press, 1998)

----- . "Cognitivism, A Priori Deduction, and Moore," *Ethics* 113, 3 (Apr. 2003): 557-75.

Moore, G. E. *Principia Ethica*, [1903] Revised Ed., with the Preface to the 2<sup>nd</sup> Edition and other papers, ed. and with an introduction by Thomas Baldwin (Cambridge: Cambridge University Press, 1993).

----- . "The Nature of Moral Philosophy," in *Philosophical Studies* (London: Routledge & Kegan Paul, 1922): 310-39.

@

Pigden, C. R. "Desiring To Desire: Russell, Lewis and G. E Moore." Pp. 244-260 in @.

Perry, R. B. *General Theory of Value* (Cambridge, MA: Harvard University Press, 1926).

Reprinted pp. 138-151 in Sellars and Hospers.

Putnam, H. *Reason, Truth and History* (Cambridge: Cambridge University Press, 1981: 205-211).

Ridge, M. "Moral Non-Naturalism," *The Stanford Encyclopedia of Philosophy* (Summer 2006), Edward N. Zalta ed.

Schilpp, P. A. *The Philosophy of G. E. Moore* (La Salle, IL: Open Court, 1942).

Sellars W. and J. Hospers *Readings in Ethical Theory*, 2<sup>nd</sup>. Ed. (New York: Appleton-Century Crofts, 1970).

Shafer-Landau, R. "Ethics as Philosophy: A Defense of Ethical Nonnaturalism," in *Metaethics After Moore*, Terry Horgan and Mark Timmons, eds. (Oxford: Clarendon Press, 2006): 209-232.

Sharp, F. C. *Ethics* (New York: The Century, 1928).

- Shaver, R., "Non-naturalism." Pp. 283-306 in @.
- Smith, M. "Moral Realism." Pp. 15-37 in *The Blackwell Guide to Ethical Theory*, Hugh LaFollette, ed. (Oxford: Blackwell, 2000).
- Snare, F. 'The Open Question as Linguistic Test,' *Ratio* 17 (1975): 123-29.
- Stratton-Lake, P. "Introduction," pp. 1-28 in *Ethical Intuitionism: Re-evaluations*, ed. by P Stratton-Lake (Oxford: Clarendon Press, 2002).
- Sylvester, R. *The Moral Philosophy of G. E. Moore* (Philadelphia, PA: Temple University Press, 1990).
- Warnock, M. *Ethics Since 1900* (London: Oxford University Press, 1960).
- West, H. "Mill's 'Proof' of the Principle of Utility," in *Mill's Utilitarianism: Critical Essays*, D. Lyons, ed. (Lanham, MD: Rowman & Littlefield, 1997): 85-98.
- White, A. *G. E. Moore: A Critical Exposition* (Oxford: Blackwell, 1958).
- Williams, B. *Ethics and the Limits of Philosophy* (Cambridge, MA: Harvard University Press, 1985).
- Wright, C. "The Perils of Dogmatism." Pp. 25-48 in @.