

Naturalism in Metaethics

Jussi Suikkanen

Final Author Copy: Published in *Blackwell Companion to Naturalism*, Kelly James Clark (ed.), Wiley-Blackwell, 2016.

This chapter offers an introduction to naturalist views in contemporary metaethics.¹ Such views attempt to find a place for normative properties (such as goodness and rightness) in the concrete physical world as it is understood by both science and common sense. The chapter begins by introducing simple naturalist conceptual analyses of normative terms. It then explains how these analyses were rejected in the beginning of the 20th century due to G.E. Moore's influential open question argument.

After this, the chapter considers what good general reasons there are for defending naturalism in metaethics. The bulk of the chapter will then survey new semantic and metaphysical forms of naturalism, which in different ways attempt to address Moore's objection. These more recent versions of naturalism – using new resources from philosophy of language, philosophy of mind, metaphysics, philosophy of science, and epistemology – attempt to explain why the open question argument fails.

Naturalist Conceptual Analyses of Normative Terms

The paradigmatic method of philosophical investigation has traditionally been conceptual analysis.² When philosophers use this method, they are attempting to capture in simpler terms what a person means when she uses a more complicated term. It is natural to think that what a

¹ Metaethics is the study of the nature of normative judgments and normative properties and the meaning of normative language. This is to be contrasted with normative ethics, which studies questions such as which actions are right and wrong.

² For different ways to understand conceptual analysis and a detailed historical overview of its use throughout the history of philosophy, see Beaney (2014).

person means when she says that “Tom is a bachelor” is that Tom is both male and unmarried. Likewise, by calling an animal “a vixen,” you mean that it is female and a fox. Conceptual analyses thus attempt to capture the meaning of more complicated terms with definitions that unpack their meaning into simpler terms. These analyses claim, for instance, that a bachelor is by definition an unmarried male and a vixen is by definition a female fox.

Philosophers tend to be interested in terms that are more central parts of our language than “bachelor” and “vixen.” One essential part of language which philosophers are interested in includes moral, evaluative, and normative terms. This is the area of language that we use when we discuss what is good and bad; which actions are right and wrong; what we ought to do; what we have reasons to do; what is just, courageous, evil, beautiful, and so on. We also use these terms in our own practical deliberations that lead us to act.

Historically, there have been many attempts to offer conceptual analyses of the central normative terms. It has been suggested, for example, that if you say that “It is good to have a job,” what you really mean is that you desire to have a job;³ this suggested conceptual analysis of the term “good” unpacks the meaning of this analyzed term in terms of the simple idea of what you desire. Likewise, it has been suggested that if I say that an act is “right” what I really mean in simpler terms is that the act brings about more pleasure for everyone than anything else that I could do.⁴ According to this analysis, the correct definition of the term “right,” which captures its meaning, is “What brings about more pleasure for everyone than anything else an agent could do in her situation.”

These simple analyses of “good” and “right” try to explain what these terms mean by talking about only ordinary, easily understandable things like desires and pleasure. Since there is nothing mysterious about human desires or the experience of pleasure, these analyses also explain how we could come to know what is good or right. These analyses furthermore say nothing about God or God’s plans. In this sense, it is easy to call these analyses naturalist analyses of “good” and “right,” in contrast to supernatural analyses, which would make sense of the meaning of “good” and “right” in terms of, say, God’s will.

³ Thomas Hobbes is often interpreted to have endorsed this definition in *Leviathan* (see Hobbes 1651/1994, I.6). Whether this was his intention remains controversial.

⁴ This analysis is often attributed to John Stuart Mill (1871/1998, 59).

Moore's Open Question Argument

Despite the attractions of naturalist explanations of the meanings of normative terms, many philosophers are suspicious of them. This line of thought is already present in the work of Richard Price (1723–1791), Ralph Cudworth (1617–1688), Francis Hutcheson (1694–1746), Thomas Reid (1710–1796), and Henry Sidgwick (1838–1900).⁵ It culminates in G.E. Moore's open question argument, which Moore formulated in his *Principia Ethica* (Moore 1903/1993, §§ 11–14). In their arguments against naturalist analyses of normative terms, Moore and the other philosophers before him rely on an intuitive standard for when an analysis successfully captures what a speaker means when she uses a given term.

To see how this test works, consider the previous example of the term “vixen.” I suggested that when a competent speaker uses this term to describe an animal, what she means is that the animal is female and a fox. Moore observed that if this definition really captures what the speaker means, then the speaker should not be able to begin to consider the question “X is a female fox, but is X a vixen?” This is a closed question for the speaker, because she both means “female fox” by the word “vixen” and has already concluded that X is a female fox. Thus, it would make no sense for her to begin to wonder whether the female fox is also a vixen. As a consequence, Moore suggested that successful analyses always lead to closed questions that have the form of the previous question.

Moore then used this test to argue that there cannot be any successful definitions of normative words in naturalist terms. Consider the two analyses previously suggested: “good” means what you desire and “right” means what brings about more pleasure than anything else that you could do. Moore argued that competent speakers can always begin to consider how to answer the following questions: “I desire X, but is X good?” and “X brings about more pleasure than any of the other alternatives, but is X right?” From this observation, he concluded that being what we desire or what brings about more pleasure than other alternatives cannot therefore be what we mean when we use the words “good” and “right.” In fact, Moore concluded from these examples that no naturalistic analyses of normative terms could ever capture what the users of

⁵ For references, see Korsgaard (2008, 307, fn. 23). See also Sidgwick (1907, bk. 1, ch. 3).

these terms mean, because the relevant questions we get from these analyses will always remain open.

Moore's argument was incredibly important for the development of modern metaethics. One reason for this is that Moore concluded from his argument that the normative terms must, as a consequence of the open questions, refer to *non-natural* properties of their own unique kind.⁶ With this, Moore single-handedly introduced the standard terminology and dividing lines for naturalism versus non-naturalism debates in metaethics. Non-naturalists defend the Moorean view according to which normative properties are in some fundamental sense of a different kind than the other, more ordinary properties that objects, actions, and states of affairs are usually thought to have. In contrast, naturalists in metaethics think that normative properties are fundamentally of the same kind as other ordinary properties.

Natural and Non-Natural Properties

After Moore's groundbreaking work, it has proven difficult to specify accurately the difference between natural and non-natural properties. There are three main ways to draw the distinction. The first way to draw it is to begin from the observation that we can use our senses to observe which objects have ordinary properties and which don't. We can see the shape of a car, hear how far a bird is from us, and feel how heavy a bag is. The suggestion, then, is that natural properties are all those properties that are empirically observable in this way, whereas whether something has a non-natural property must be known in some other way (Copp 2003).

The second holds that ordinary properties are a part of the causal network of the world. For example, the steepness of a hill causes a ball to roll down and the speed of a bullet causes a lot of damage inside a human body. The claim, then, is that natural properties are all those properties that are a part of these sorts of causal chain. The having of these properties can both be a consequence of other events that take place in the real world and it can also cause other things to happen. A defining feature of non-natural properties is that they are in some sense outside the causal network of the physical world (Lewis 1983, § 2).

⁶ For an insightful investigation of Moore's conclusion, see Dreier (2006).

The third is to argue that natural properties are special because scientists rely on these properties in the best systematic explanations of what happens in the world (Little 1994; Shafer-Landau 2006, 211). As a consequence, because physics relies on properties such as spin and charge in its most fundamental explanations, these properties are natural properties. The claim, then, is that the properties that are not used in systematic scientific explanations are for this reason of a different kind than natural properties. Such properties are non-natural properties.

These three ways of distinguishing natural properties from non-natural properties have their own problems. They do suggest, however, that at the heart of the disagreement between naturalists and non-naturalists is the question of whether morality is autonomous from empirical and scientific investigation.

I have already explained why G.E. Moore believed that the meaning of normative terms cannot be analyzed in simpler terms which refer to ordinary natural properties. I also mentioned that Moore concluded from the open question argument that unanalyzable normative terms must therefore refer to simple non-natural properties of their own unique kind. Most developments in metaethics during the 20th and the early 21st centuries can be understood as reactions to Moore's open question argument. The advantages of naturalism provided a powerful motive for developing innovative and sophisticated naturalist theories of normative terms and properties that could address Moore's objection to naturalism. Before I introduce these new forms of naturalism, I will first explain the advantages.

General Arguments for Naturalism in Metaethics

There are at least four good reasons for preferring naturalism⁷ in metaethics over the Moorean alternatives that are committed to the existence of additional non-natural properties. The strength of these reasons explains in part the ingenuity of the naturalist responses to Moore.

First of all, metaphysical parsimony is an attractive methodological principle in all scientific theorizing.⁸ The basic idea of this principle is that we should always look for the

⁷ By "naturalism," I mean substantive naturalism: a view about normative concepts and properties. Methodological naturalism, in contrast, is the view that philosophy (ethics included) should proceed by *a posteriori* empirical investigation. See Railton (1989, 155–157).

simplest theoretical explanations of the events we observe: the kinds of explanation that rely on the fewest possible objects, properties, and laws. Simple explanations have proven to be incredibly powerful, which is why the principle of parsimony has such a good track record. The simple explanations of science help us to predict successfully what happens across a wide variety of different contexts. This has led to the rejection of our reliance on spirits or magic in our explanations and predictions, and it is why we have made genuine intellectual progress.

Because the goal of aiming at metaphysical parsimony has served science so well, we should also rely on this principle in other domains of rational inquiry (including philosophy). Unless it is necessary, we should also avoid positing new kinds of property to explain reality. This is a powerful reason for not adding any non-natural properties to our worldview.

The second reason for preferring naturalism has to do with how we can know about normative properties.⁹ As already explained, naturalists believe that normative properties are exactly like other ordinary properties. When it comes to those properties, there is a relatively simple and unproblematic explanation for how we know about them: centrally, ordinary natural properties causally impact our senses. I can know that the book on my desk is rectangular because the photons it reflects hit my retina, which is sensitive to these particles. Through this causal process, I come to see the shape of the book, and through this perception, I reliably come to know that the book is rectangular.

Because naturalists believe that normative properties are like other ordinary properties, they claim that we can know which objects, actions, and states of affairs have these properties by the same empirical methods. For naturalists, normative properties are a part of the causal network of the world, and for this reason naturalists claim that we can observe normative facts as we do all other facts.

If you believe that normative properties are non-natural, however, it will be difficult for you to explain how we come to know about them. Nonnatural properties (1) cannot be observed,

⁸ This principle is sometimes called Occam's razor. For an overview of its role and its attractions in philosophical and scientific theorizing, see Baker (2010).

⁹ For *locus classicus*, see Mackie (1977, 38). For a more recent powerful development of the argument, see (Bedke 2009).

(2) are not part of the causal nexus of the world, and (3) are not used in the best scientific explanations. These features of non-natural properties rule out using both our senses and science in knowing when an act is good or right. Non-naturalists must defend *a priori* methods in the epistemology of normative properties. While they often talk about self-evident normative principles that we discover through pure *a priori* reasoning, it is unclear how such reasoning is sensitive to the non-natural properties that non-naturalists believe are part of the furniture of the world.¹⁰

Third, it is equally difficult for non-naturalists to explain how normative terms came to refer to non-natural properties. In most cases, our words refer to certain objects and substances at least in part because we causally interact with them (Putnam 1975; Kripke 1980). We drink and bathe in a certain liquid, which at some point in human history we named “water.” This causal connection explains how the word “water” came to refer to the substance it refers to. In the same way, a certain baby was named “John F. Kennedy.” Through the interaction between that baby and the people who gave him his name, that name came to have its reference. The names of substances and individuals, then, refer to specific substances and individuals (rather than to some other substances and individuals) because we are in a causal contact with them when we introduce the names.

If we cannot causally interact with alleged non-natural normative properties, then there will have to be some other explanation for how the word “right” came to refer to the property of rightness (rather than to some other property). With the exception of Ralph Wedgwood’s recent work (2007, ch. 4), however, non-naturalists have failed to explain how words come to refer to the relevant non-natural properties. An important advantage of naturalist views in metaethics is that their reliance on causal connections ensures that normative terms refer to the right properties.

¹⁰ See Bedke (2009). Devitt (2010) develops general arguments against *a priori* knowledge. For non-naturalist responses to these worries, see, for example, Audi (2005), Shafer-Landau (2003, pt. 5), Wedgwood (2007, pt. 3), and Enoch (2011, ch. 7).

The final advantage of naturalism is the failure of non-naturalists to explain how normative properties supervene on natural properties.¹¹ In order to unpack this abstract idea, we can begin from the thought that it seems obvious that if two actions are identical in all their ordinary natural features, they are also the same normatively speaking: equally good and right. However, if normative properties were distinct and additional properties of their own unique kind, it would be difficult to explain why different instances of otherwise identical actions could not have different normative properties. Naturalists, in contrast, have again an easy answer to this challenge. They, after all, claim that normative properties belong to the set of ordinary natural properties, and therefore it is trivially true that there is no difference in the normative properties that two actions have without a difference in their ordinary natural properties.

To summarize, there are a number of powerful reasons for accepting a naturalistic theory of normative terms and properties. The strong intuitions behind Moore's open question argument also give us an idea of what a satisfactory naturalistic view in metaethics would look like. It would need to explain, for example, why it seems that when we talk about what is good or right, we are not merely talking about what we desire or what brings about the most pleasure. How different naturalist views have reacted to this intuition helps us to classify the new forms of naturalism in metaethics. We will first look at naturalist responses to Moore that focus on normative terms. Following this, we will consider views that focus on the nature of the normative properties.

Semantic Forms of Naturalism in Metaethics

In this section, I will introduce different forms of naturalism which attempt both (1) to capture the meaning of normative language in broadly naturalist terms and (2) to address Moore's concern that such analyses fail because they lead to open questions. I will first outline the expressivist response to Moore's open question argument. I will then explain how simple

¹¹ This objection was first made by Simon Blackburn (1993a,b). For a non-naturalist response, see Shafer-Landau (2003, 84–89). Frank Jackson has argued that supervenience itself collapses any additional alleged non-natural properties to the base-level natural properties (1998, 122–123). For a critical investigation of this argument, see Suikkanen (2010).

reductive naturalist analyses of normative terms could be defended against Moore's argument by taking into account pragmatic considerations that influence our willingness to use these terms. Finally, I will outline Frank Jackson's network analyses of normative terms.

Expressivism as a Version of Naturalism in Metaethics

Soon after the publication of *Principia Ethica*, philosophers started to consider why we have the intuition that, whatever natural properties an act has, it still remains an open question whether that act is right. One insightful answer to this question came from philosophers who defended early versions of expressivism (Ayer 1936, 106–110; Hare 1952, 94–110).

Expressivists hold that there is a necessary connection between thinking that an act is good and desiring to do that act.¹² Many people intuitively believe that you cannot sincerely conclude that it is good to help other people unless you also have at least some desire to do so. In order to explain this internal connection between normative judgments and motivation, expressivists believe, roughly, that thinking that an act is good consists itself in a desire to do that act (rather than believing that the act has a normative property).¹³ Expressivists also claim that the meaning of the sentence "X is good" can be analyzed wholly in terms of the desire-like attitudes for doing X that this sentence conventionally expresses.

Expressivists use this theory to provide an illuminating explanation of why Moorean questions always remain open. When I judge that an act brings about more pleasure than its alternatives, I form a belief of a natural property that the act in question has: I believe that this act has the property of bringing about a lot of pleasure. According to expressivists, judging that an act is good also requires having a desire to do that act. However, no belief about what natural properties an act has can itself settle what I desire. Knowing that it is sunny outside does not move me to go outside, unless I want to be in the sun. This desire is always something distinct from knowing that it is sunny. Forming a desire to do something is therefore something

¹² This thesis is usually called "judgment internalism." For a clear overview of the debates concerning this thesis, see Björklund et al. (2012).

¹³ For representative recent expressivist texts, see Blackburn (1998), Gibbard (2003), and Schroeder (2008).

additional that I must always do after I have formed the relevant beliefs about the situation I am in.

According to expressivists, the additional motivation required when using normative terms explains why questions of the form “X is N, but is it good?” always remain open, no matter what natural property N is. When you ask that form of question, you are considering whether you want to do X when you know that X has the property of being N. Whatever natural property N is, thinking about this always makes sense, which is why Moorean questions will always remain open.

Expressivism is a naturalist view, and therefore it shares the general advantages of naturalism (Blackburn 1984; 1998, 49).¹⁴ Expressivists give a fully naturalist explanation of which mental states constitute normative judgments, and this account does not require the existence of any non-natural properties. In fact, all that is required to make sense of normative judgments, according to expressivists, is ordinary, desire-like attitudes, which are commonly recognized in the human sciences, such as psychology and economics.

When metaethicists discuss naturalism, they typically have in mind cognitivist views, according to which normative judgments are ordinary beliefs about natural properties. Naturalism in metaethics, then, is assumed to be a form of cognitivism. But cognitivism is in conflict with noncognitivist expressivism. While expressivists endorse a fully naturalist worldview, they reject the typical cognitivist assumption.

So far, I have only explained what expressivists think about the nature of normative judgments; I have not said anything about what they think about normative properties. The reason for this is both (1) that what normative properties are like is less important from the expressivist perspective and (2) that expressivists disagree about this issue. Some expressivists think that normative properties are metaphysically lightweight “projections” of our desire-like attitudes (Blackburn 1993c, 55–60). While these “shadow properties” are in one sense of their own unique kind (and so “non-natural”), they lack the kind of ontological weight which non-naturalist realists assume that normative properties have.

¹⁴ In the same way, it should be emphasized that moral error theory too is a form of naturalism. On this view, semantically normative claims presuppose the existence of non-natural properties, but, metaphysically speaking, there are no such properties which make the normative claims false (see Mackie 1977).

The standard way to illustrate this view is by analogy to the property of being disgusting. In the order of explanation, the reaction of being disgusted by different objects comes first. When we talk about disgusting things, we project our reaction of disgust on to those things. In the same way, many expressivists think that we react to the world by forming different desire-like attitudes. When we talk about normative properties, we project these reactions on to different actions and states of affairs in the same way as in the case of disgusting things. As a consequence, normative properties are shadows of our desire-like attitudes.

Other expressivists think that normative properties are ordinary natural properties (Gibbard 2003, ch. 5).¹⁵ We can again illustrate this view with the example of disgusting things. As a consequence of projecting our reactions of disgust on to worldly objects consistently, there will be a complex natural property that is coextensive with things deemed disgusting. This can be the property of having the property P or the property Q or the property R and so on, where the predicates “P,” “Q,” “R,” and so on describe in detail the natural properties of each and every instance of a disgusting thing. So, roughly, it’s the property of being a rotten egg, or a case of incest, or similar. One view of the property of being disgusting, then, is that it is constituted by the underlying coextensive natural property, which can take a form of infinitely long disjunction.

Some expressivists believe that, in the same way, normative terms refer to an underlying, perhaps disjunctive natural property that is coextensive with the set of good things or right things. The property of being right can, on this expressivist view, be roughly the property of being a case of helping an old person across the street, or giving money to charity, or not stealing from someone, and so on. It is just that when we talk about this natural property by using normative terms such as “right,” we are expressing our desire toward doing actions that have this natural property. This is something we could not do by merely using ordinary non-normative descriptive terms.

¹⁵ Gibbard’s argument that there is no logical room over and above the base-level natural properties is similar to Jackson’s. See Gibbard (2003, 94–98) and Jackson (1998, 122–123).

Implicatures and Know-How

In this subsection, I will introduce two attempts to defend simple naturalist reductive analyses of normative terms with the help of recent developments in the philosophy of language. The first view is based on the distinction between what a word means literally (semantics) and what can be done in practice by using that word (pragmatics). The second view is based on the distinction between propositional knowledge and skills.

To illustrate the first distinction, we can use Paul Grice's classic example of a linguistic phenomenon he calls "a conventional implicature" (1989, 46). Consider the following sentences:

- (1) Alice is poor but she is honest
- (2) Alice is poor and she is honest

The first observation we should make about these two sentences is that they are true in exactly the same situations: both are true when Alice is both poor and honest (it is attractive to understand what a sentence means in terms of the conditions in which it is true).¹⁶ If we accept this simple idea, then we are led to think that the literal meanings of both (1) and (2) are exactly the same.

Despite this, we would not want to utter these sentences in exactly the same situations. The reason for this is that (1) communicates a further thought which (2) doesn't. In practice, the use of "but" in (1) suggests that poor people are not usually honest. The sentence does not say this literally, but if you were to use it in discussion you would communicate this thought to your audience. This example suggests that many words and utterances imply things that are not part of their literal meaning.

¹⁶ Truth-conditional semantics culminate in the work of Donald Davidson. See the essays collected in Davidson (1984).

It is possible to use this idea to defend simple naturalist analyses of normative terms against Moore.¹⁷ Let us assume that expressivists are correct about why we think that questions of the form “X is N, but is X good?” are open questions: because uttering the word “good” commits you to desiring X, whereas thinking that X is N does not commit you to this desire, whatever natural property N is. If we accept that words and sentences can imply things that are not part of their literal meaning, then naturalists can argue that using the word “good” commits you to having certain desires because using this word implies that you have them (and not because the relevant desires have anything to do with the literal meaning of the word).

To see how this works, let us return to the claim that “right” means whatever brings about the most pleasure. Why is the question “X brings about most pleasure, but is X right?” an open question? The naturalist response to Moore is that merely describing an act as the one that brings about most pleasure does not commit you to desiring anything, whereas saying that this act is right implies that you want to do X. Accordingly, the utterance that “X is right” would have this implication even if, in saying that X is right, one literally meant that X brings about most pleasure. This then explains why you can hesitate to call an act that brings about most pleasure right even if by saying that it is you mean that it brings about the most pleasure. Saying that the act is right, despite its literal meaning, implies something about your own motivations.

This response to the open question argument leaves the naturalist with one important task. There are many candidates for a naturalist analysis of normative words. “Right” could mean “whatever brings about most pleasure” or “what society accepts” or any number of other things. Naturalists owe us a plausible theory of which one of these analyses we should accept.

One problem is that the practical implicature response to the open question argument means that what we would say about individual cases is not a reliable indication of what we literally mean when we use normative terms. After all, our willingness to use these terms is also influenced by the implicatures these terms carry with them in concrete situations. This is why naturalists must give us an account of how we are to determine what these words mean without relying on our intuitions about individual cases. One suggestion is that we should look at general

¹⁷ Proposals along these lines have been defended, for example, by David Copp (2001) and Stephen Finlay (2005). Copp believes that normative utterances conventionally implicate desire-like attitudes, whereas according to Finlay the implicatures in question are conversational implicatures.

empirical semantic theories investigated in linguistics. These theories might be able to tell us what literal meanings normative terms must have in order to explain how these terms combine with other linguistic expressions to form meaningful sentences across different contexts.¹⁸

There is also a second way in which relatively simple naturalist analyses of normative terms have been defended against Moore's open question argument. This second proposal begins from the idea that we do not learn new terms in the form of explicit definitions when we grow up (Smith 1994, 37–38). For example, when I learned to use the word "car" correctly, I did not learn this word as a definition of the form "X is a car if and only if X has four wheels, an engine,..." Rather, I become a competent user of this term by acquiring the skill of classifying objects into two categories (those that are cars and those that are not). When I rely on this skill as a speaker, I needn't have any explicit definition in mind, even if I do make use of some implicit standards that constitute my conceptual competence with the word in question.

This idea can be used to construct an attractive response to the open question argument. It can explain why a speaker can begin to consider the question "X is N, but is X good?" even when what she means by saying "X is good" is that X is N (Smith 1994, 36–38). When a speaker utters the sentence "X is good," she is relying on the skill of being able to classify things according to certain implicit standards, which she learned in the form of know-how when she was growing up. What these standards are will not be transparent for the speaker. It can turn out that she is classifying things as "good" by relying implicitly on the standard "Is X N?" In this case, by calling things "good," she would mean that these things are N, even if she might not be aware of this. Likewise, by calling an act "right," a speaker may well be categorizing it as one that brings about the most pleasure even if she is not aware that she is using this standard. This is why Moore's questions of the form "X is N, but is X good?" can remain open even if the definitions from which they are formed correctly capture what the speaker implicitly means.

If naturalists give this response to the open question argument, they still need to explain how we can discover what criteria speakers are implicitly using to classify things, and thus what constitutes the meaning of the terms they use. One suggestion is that, if we let speakers reflect

¹⁸ On the basis of this type of consideration, Stephen Finlay has suggested that the literal meaning of "good" in a given sentence can, for example, be analyzed in terms of the promotiveness of an end which is specified by the context of utterance (2014, ch. 2).

carefully enough on what role the analyzed terms play in their lives and how they use them, they should be able to discover what standards they are relying on implicitly when they use those terms (Smith 1994, 38).

Michael Smith has used this idea to construct a naturalist theory of the meaning of normative terms. According to Smith, one fundamental part of the role of normative terms is that making judgments with these terms creates requirements of rationality (1994, ch. 3). Roughly, on Smith's view, if you think that doing an act is good, then either you desire to do that act or you are being practically irrational.¹⁹ Smith then thinks that when we analyze normative terms like "good," we should find definitions that help us to explain why these terms create the previous type requirements of rationality. Smith concludes that when you say that we have a normative reason to do a certain act in a certain context, what you mean is that more rational (more informed, unified, and coherent) versions of us would want us to do that act in that context (1994, ch. 5). This naturalist analysis of the term "normative reason" might then explain the role which this term has in our deliberations, and might also be something that we can identify as the implicit standard we use in considering what we have reason to do.

Network Analysis

Frank Jackson has argued that, even if we cannot give simple naturalist analyses of normative terms, we can still reductively analyze normative language as a whole in a naturalist fashion (1998, ch. 5). According to this view, we will not be able to find naturalist definitions of the form "To be good is by definition to be N," where "N" is something simple like "what we desire" or "what brings about most pleasure." Jackson argues that, despite this, we can capture the whole network of normative terms in naturalist terms.

Jackson offers a simple recipe for how the more sophisticated naturalist network analyses of normative terms can be constructed. According to Jackson, normative terms form an interrelated network of concepts. The place a given normative term has in this network determines both (1) the meaning of that term and (2) what a competent speaker must understand about that term (Jackson 1998, 129–130).

¹⁹ This thesis is a form of conditional judgment internalism (Björklund et al. 2012, § 2).

At the first stage of Jackson's analysis, we collect all platitudes that, for example, have to do with "right." These trivial claims connect this term to other normative concepts. They might include claims such as:

- If someone judges that X is right, she is disposed to do X.
- If one person says that X is right and another person that it isn't, then at most one of them is correct.
- Right actions in some sense express equal concern and respect to everyone.
- Right actions tend to have good consequences.
- ... (Jackson 1998, 130–131; see also Smith 1994, 39–41)

From this, we can form a very long conjunction of all the platitudes that connect "right" to other normative and non-normative terms. This conjunction will look like this:

1. (If someone judges that X is right, she is disposed to do X) and (if one person says that X is right and another person that it isn't, then at most one of them is correct) and (right actions in some sense express equal concern and respect to everyone) and (right actions tend to have good consequences) and... (Jackson 1998, 140)

After this, we can replace all instances of normative property words in (1) with variables. As a consequence, we get:

2. (x, y, z, \dots) (If someone judges that X is x , she is disposed to do X) and (if one person says that X is x and another person that it isn't, then at most one of them is correct) and (x actions in some sense express y to everyone) and (x actions tend to have z consequences) and...

If we in addition claim that properties x, y, z, \dots specified by (2) exist and that they exist uniquely, we get the following:²⁰

3. There are (x, y, z, \dots) and for all (x^*, y^*, z^*, \dots) [(If someone judges that X is x , she is disposed to do X) and (if one person says that X is x and another person that it isn't, then

²⁰ This is called a Ramsey sentence, because Frank Ramsey relied on similar sentences in an attempt to represent the empirical content of scientific theories (Ramsey 1978).

at most one of them is correct) and (x actions in some sense express y to everyone) and...] and [(if someone judges that X is x^* , she is disposed to do X) and (if one person says that X is x^* and another person that it isn't, then at most one of them is correct) and (x^* actions in some sense express y^* to everyone) and...] if and only if (x is identical with x^* and y is identical with y^* and...). (Jackson 1998, 140)

This sentence is formulated in purely naturalist terms. It can be used to offer an *a priori* reductivist naturalist analysis of the term “right.” According to this analysis, rightness is by definition the unique natural property that is related to other properties exactly in the way that the big conjunction of platitudes tells us that normative properties are related to one another (Smith 1994, 45–46).

The thought behind this complicated analysis is simple. The role that the word “right” plays in our ordinary normative discourse is like a symbol on a map. A map picks out the place of a specific location by specifying how the location is related to all the other locations shown. In the same way, this analysis of the term “right” picks out a certain natural property by specifying how that property is related to other properties in ordinary language and thinking.

This form of naturalism thus enables us to define normative terms as terms that refer to the unique natural properties (whatever they may be) that play the role defined by the place of those terms in the whole network of normative concepts. It is then a matter of further empirical research to discover which particular natural properties the concepts in the whole network of normative concepts refer to. It is true that we could form a relevant open question from the previous definition of “right” which speakers could begin to consider. However, Jackson argues that this definition will only create open questions, because speakers can always be hesitant about whether the discovered analysis based on the network of platitudes is correct. Given how complicated the network analysis is, it is no wonder that it will not be apparent for the speakers that this definition correctly captures what “right” means.²¹

²¹ One objection to Jackson’s moral functionalism is that the network of platitudes in the case of normative concepts is not rich enough to pick out unique natural properties as the referents of these terms. For this permutation problem, see Smith (1994, 48–54).

Synthetic Forms of Naturalism in Metaethics

All the naturalist views introduced so far have focused on normative terms and language. They all attempt to show that, in one way or another, normative terms can be defined in non-normative language that refers only to ordinary natural properties. However, there is another alternative. Many metaethicists have recently attempted to naturalize normative properties directly on the level of metaphysics without offering naturalist semantic analyses of normative terms.²² This new project was made possible by new developments in metaphysics, philosophy of science, and philosophy of language during the second half of the 20th century.

The so-called synthetic naturalists in metaethics think that even if normative terms cannot be defined in simpler naturalist terms, normative properties can still be natural properties. Nonreductive synthetic naturalists believe that normative properties are natural properties of their own, whereas reductive synthetic naturalists think that normative properties are reducible to other natural properties.²³

We can start approaching these views from the idea that, in a sense, the new synthetic forms of naturalism accept the first conclusion of Moore's open question argument (there are no correct definitions of normative terms) but reject the second one (therefore, these terms refer to non-natural properties). When Moore was writing, this combination was not a real option, because it was thought that the meaning of a word determined its reference. If you accept this principle of Frege's philosophy of language (1892/1997, 152), then it is tempting to assume that, if two terms differ in meaning, they must refer to different properties. However, Hilary Putnam (1975) and Saul Kripke (1980) famously argued that there is no reason to accept Frege's principle. They showed that there can be two terms that have different meanings which nonetheless refer to the same individual or substance in all possible worlds. It is easy to use scientific identity statements to illustrate this idea.²⁴

²² These philosophers include Richard Boyd (1988), Nicholas Sturgeon (1988), David Brink (1989), and Mark Schroeder (2007, ch. 4). Because the first philosophers defending the theory worked at Cornell at the time, synthetic forms of naturalism are often called "Cornell realism" in metaethics.

²³ Sturgeon (1988, 239–241) is an example of a nonreductivist synthetic naturalist, whereas Schroeder (2007, 61) is a clear example of a reductivist synthetic naturalist.

²⁴ Naturalists in metaethics who use this scientific identity analogy include Boyd (1988, 196) and Brink (1989, 157).

According to scientists, water is H₂O molecules. Despite this metaphysical identity, “water” and “H₂O” have different meanings. When you explain the meaning of “water,” you will say things like “Water is a transparent, tasteless liquid,” “It falls from the sky and fills lakes and rivers,” and so on. In contrast, the meaning of the theoretical term “H₂O” has something to do with how hydrogen and oxygen atoms can form larger molecules. Kripke and Putnam’s crucial observation was that not everything that is a sample of a transparent, tasteless liquid that falls from the sky and fills lakes and rivers will be water. We can imagine a planet where we find a substance that is like this but is not H₂O (but rather has some other chemical constitution, XYZ). The samples of this substance will intuitively not be water, because we only count H₂O as water. This means that here we have a case where two terms that differ in meaning refer to the very same substance in all worlds.

This analogy creates logical room for the new synthetic forms of naturalism in metaethics. These views hold that even if the predicates “good” and “N” differ in meaning, they can still refer to exactly the same natural property N across all worlds.²⁵ That is, goodness can be identical with the property of being N even if this identity is not based on the meanings of the relevant words. As a result of this possibility, Moore can be right that the meaning of normative words cannot be captured in simpler terms and yet wrong about normative properties being distinct non-natural properties. This is why the traditional open question argument cannot be used as an objection to the new synthetic forms of naturalism.

All of this, of course, leaves synthetic naturalists with the question: Which natural properties exactly do constitute normative properties metaphysically? According to Cornell realists, in the same way that we use the best overall science of our day to determine which substance is water, we should also rely on the best overall ethical theory to determine which natural properties constitute normative properties. If the simple utilitarian ethical theory (right actions maximize the amount of pleasure) turns out to be true according to the best theory, then the natural property of maximizing pleasure constitutes the property of rightness.

²⁵ This view is most clearly defended by Boyd (1988), Brink (1989), and Schroeder (2007). Peter Railton (1986) has suggested that once we find a naturalist property that plays the role specified by an ordinary normative term, we can revise the meaning of that term so that it will mean the same as the non-normative terms that pick out that property.

How do we know which ethical theory is the best? At this point, Cornell realists are inspired by the epistemological debates in philosophy of science. Quine argued convincingly that individual elements of scientific theories or even complete theories cannot be empirically tested by observations (1969, 79). What the observations show will always depend on many background assumptions outside them. This is because, by making corrections somewhere else in your web of beliefs, you can always make your theory fit what you observe. For this reason, Quine thought that we can only test worldviews against observations holistically (in their entirety).

The new forms of synthetic naturalism in metaethics borrow this form of coherentism from the philosophy of science.²⁶ They state that we should first attempt to seek a wide reflective equilibrium between our intuitions about cases, general moral principles, and, more broadly, common-sense beliefs and scientific theories about the world. The best overall theory we arrive at as a result of this procedure will then tell us which natural properties constitute normative properties.

Finally, the development of the previous type of synthetic naturalism in metaethics led to an important debate about the nature of normative properties. I have already suggested that one important criterion for whether a given property is a natural property is whether the property can play a role in systematic causal explanations. Some people additionally think that we should believe in the existence of only those properties that are indispensable in causal explanations. Gilbert Harman argued against the new forms of naturalism in metaethics by relying on this principle and the claim that normative properties are not needed in the best causal explanations (1977, ch. 1).

Harman compares two cases. In one, a scientist sees a vapor-trail in a cloud chamber and, on the basis of this observation and her background theoretical beliefs, infers, “There goes a proton.” Harman claims that, in this case, the best explanation of the scientist’s observation must

²⁶ See, for example, Boyd (1988, 191–192) and Brink (1989, ch. 5). Some metaphysicians consider this type commitment to Quinean confirmation-holism to be a defining feature of naturalism (Devitt 2010, 274). The coherentist methodology also fits nicely with the view that normative terms refer to the natural properties that best explain our use of those terms (Boyd 1988, 195).

refer to the proton itself, which then is just one reason why we should believe that the proton exists. Without the proton, there would not have been an observation.

In the second case, a group of hoodlums sets a cat on fire, which leads a passer-by to judge, “That is wrong!” As in the previous case, without her background ethical theory of which actions are wrong, the passer-by would not have made this observation. However, Harman argues that there is no need to refer to the property of wrongness at all in the best causal explanation of the moral judgment. All we need to refer to in that explanation is how the passer-by’s upbringing has made her disapprove of burning cats. And, because we do not need to refer to the wrongness of the act to explain the observation, according to Harman, there is no reason to believe that there is such a property in the first place.

Nicholas Sturgeon’s (1988) response to this challenge played a big role in the development of new forms of naturalism in metaethics. Sturgeon begins with the idea that the challenge creates an interesting objection to the existence of normative properties only if it shows that those properties would not play a causal role even if they existed (or otherwise the objection would beg the question). This is why, when we think about the cat-burning case, we are allowed to assume that there are normative properties. Sturgeon then argues that, if we are allowed to assume that there are normative properties, these can be used in informative causal explanations of our observations and other events.

One example of a good moral explanation is the following: Hitler ordered the deaths of millions of Jews because he was morally depraved (Sturgeon 1988, 234). In this case, the fact of Hitler’s depraved nature would be explanatorily irrelevant only if he would have ordered the deaths of millions of Jews even in the counterfactual situation in which he was not morally depraved. According to Sturgeon, the naturalist view Harman is objecting to is based on the idea that Hitler’s moral depravity is constituted by his non-normative natural properties: by his racism, insensitivity to other people’s suffering, and so on (1988, 245–246). This means that, if Hitler had not been morally depraved, he would have been a very different kind of person with respect to his non-normative natural properties. In this situation, he probably would not have ordered the deaths of millions of Jews, and therefore it seems like we can use normative

properties in informative causal explanations. This is a good reason to believe that these properties do exist and that they are natural properties.²⁷

Sturgeon's naturalist defense of normative properties against Harman's explanatory challenge led to a fruitful debate about moral explanations. Harman (1986) responded that the counterfactual Sturgeon used to test the relevant explanations was too permissive. It would mean, for example, that there was no logical room for moral epiphenomenalism: the view that Hitler's non-normative character traits explain both his actions and why he was morally depraved even if his morally depraved nature doesn't explain anything. After all, defenders of this view could not deny that, if Hitler had been morally different, he might not have done what he did. Because moral epiphenomenalism is a consistent view, Harman argues that the truth of the counterfactual "had X not been P, then Y would not have happened" is not sufficient to make property P explanatory. In response to this objection, naturalists have attempted to find other ways in which higher-order normative properties could be necessary for the best causal explanations even if they are not themselves causally efficacious.²⁸ These views attempt to vindicate the view that normative properties are real natural properties by showing that these properties are indispensable in the best causal explanations in some more sophisticated way than that Sturgeon initially assumed.

Conclusion

In this chapter, I have outlined how the general advantages of naturalist views in metaethics led to the development of new, sophisticated forms of semantic and synthetic naturalism in response to Moore's open question argument. The debates about these new forms of naturalism are one important reason why so much progress has been made in metaethics during the last 50 years or

²⁷ Sturgeon also uses this argument to motivate his nonreductivist form of naturalism. It is not necessary that the extension of the normative property that does the explanatory work (and is thus natural) can be picked out with non-normative terms that refer to other natural properties.

²⁸ The best example of an attempt to address Harman's further objection is Jackson and Pettit's (1990) notion of program explanations. For a clear overview of the whole debate about moral explanations and naturalism, see Miller (2013, ch. 8).

so. Furthermore, at some point toward the end of the 20th century, as a consequence of these developments, naturalism seems to have become the default view in Anglo-American academic metaethics.

More recently, non-naturalism has made a comeback.²⁹ There are several reasons for this. First, the new forms of naturalism outlined in this chapter continue to be critically debated, because each faces significant theoretical challenges of its own. Second, some non-naturalists continue to argue that naturalist views in metaethics have implausible first-order normative consequences. For example, the so-called particularists claim that these views cannot account for the complicated way in which whether some consideration is a reason for an act depends on its context (Dancy 2004, ch. 4). Others insist that naturalist views cannot explain why we have reason to care about our future suffering independently of what we care about currently (Parfit 2011, chs. 3 and 24–30).

The third challenge to naturalism is based on the moral twin-earth cases (Hare 1952, 146–149; Horgan and Timmons 1990). This objection is based on the idea that naturalists must describe a causal mechanism that fixes the reference of normative terms to the relevant natural properties. We can then imagine cases in which this mechanism fixes the reference of a normative term to different properties in different communities. If naturalism were true, it would follow from this that these communities would thereafter lack shared concepts with which to have normative disagreements. Some people find this awkward consequence a good enough reason to reject naturalism. Finally, it is also argued that, perhaps with the exception of expressivism, the new forms of naturalism are still unable to explain the close connection between normative thinking and motivation (Wedgwood 2007, ch. 3). The debate between naturalists and non-naturalists is therefore unlikely to go away any time soon.

References

Audi, R. (2005). *The Good in the Right: A Theory of Intuition and Intrinsic Value*. Princeton: Princeton University Press.

²⁹ Currently prominent non-naturalists include Russ Shafer-Landau (2003), Ralph Wedgwood (2007), David Enoch (2011), and Derek Parfit (2011).

- Ayer, A.J. (1936). *Language, Truth and Logic*. London: Gollancz.
- Baker, A. (2010). Simplicity. In *The Stanford Encyclopedia of Philosophy*. Available from: <http://plato.stanford.edu/archives/sum2014/entries/simplicity/> (last accessed July 15, 2015).
- Beaney, M. (2014). Analysis. In *The Stanford Encyclopedia of Philosophy*. Available from: <http://plato.stanford.edu/archives/sum2014/entries/analysis/> (last accessed July 15, 2015).
- Bedke, M. (2009). Intuitive Non-Naturalism meets Cosmic Coincidence. *Pacific Philosophical Quarterly* 90: 188–209.
- Björklund, F., Björnsson, G., Eriksson, J., Olinder, R.F., and Strandberg, C. (2012). Recent Work on Motivational Internalism. *Analysis* 72: 124–137.
- Blackburn, S. (1984). *Spreading the Word*. Oxford: Oxford University Press.
- Blackburn, S. (1993a). Moral Realism. In *Essays in Quasi-Realism*, pp. 111–129. Oxford: Oxford University Press.
- Blackburn, S. (1993b). Supervenience Revisited. In *Essays in Quasi-Realism*, pp. 111–129. Oxford: Oxford University Press.
- Blackburn, S. (1993c). Morals and Modals. In *Essays in Quasi-Realism*, pp. 52–74. Oxford: Oxford University Press.
- Blackburn, S. (1998). *Ruling Passions: A Theory of Practical Reasoning*. Oxford: Oxford University Press.
- Boyd, R. (1988). How to Be a Moral Realist. In *Essays on Moral Realism*, edited by Geoffrey Sayre-McCord, pp. 181–228. Ithaca: Cornell University Press.
- Brink, D. (1989). *Moral Realism and the Foundations of Ethics*. Cambridge: Cambridge University Press.
- Copp, D. (2001). Realist-Expressivism: A Neglected Option for Moral Realism. *Social Philosophy and Policy* 18: 1–43.
- Copp, D. (2003). Why Naturalism? *Ethical Theory and Moral Practice* 6: 179–200.
- Dancy, J. (2004). *Ethics Without Principles*. Oxford: Oxford University Press.
- Davidson, D. (1984). *Inquiries into Truth and Interpretation*. Oxford: Oxford University Press.
- Devitt, M. (2010). No Place for the A Priori. In *Putting Metaphysics First: Essays on Metaphysics and Epistemology*, pp. 271–291. Oxford: Oxford University Press.

- Dreier, J. (2006). Was Moore Moorean? In *Metaethics After Moore*, edited by Terry Horgan and Mark Timmons, pp. 191–208. Oxford: Oxford University Press.
- Enoch, D. (2011). *Taking Morality Seriously: A Defence of Robust Realism*. Oxford: Oxford University Press.
- Finlay, S. (2005). Value and Implicature. *Philosophers' Imprint* 5: 1–21.
- Finlay, S. (2014). *Confusion of Tongues: A Theory of Normative Language*. Oxford: Oxford University Press.
- Frege, G. (1892/1997). On sinn and bedeutung. In *The Frege Reader*, translated by Max Black, edited by Michael Beaney, pp. 151–171. Oxford: Blackwell.
- Gibbard, A. (2003). *Thinking How to Live*. Cambridge, MA: Harvard University Press.
- Grice, P. (1989). *Studies in the Way of Words*. Cambridge, MA: Harvard University Press.
- Hare, R. (1952). *The Language of Morals*. Oxford: Oxford University Press.
- Harman, G. (1977). *The Nature of Morality: An Introduction to Ethics*. New York: Oxford University Press.
- Harman, G. (1986). Moral Explanations of Natural Facts – Can Moral Claims be Tested Against Reality. *Southern Journal of Philosophy* 24(Suppl.): 57–68.
- Hobbes, T. (1651/1994). *Leviathan*, edited by E. Curley. Indianapolis: Hackett Publishing.
- Horgan, T. and Timmons, M. (1990). New Wave Moral Realism Meets Moral Twin Earth. *Journal of Philosophical Research* 16: 447–465.
- Jackson, F. (1998). *From Metaphysics to Ethics: A Defence of Conceptual Analysis*. Oxford: Oxford University Press.
- Jackson, F. and Pettit, P. (1990). Program Explanation: A General Perspective. *Analysis* 50: 107–117.
- Korsgaard, C. (2008). Realism and Constructivism in Twentieth-Century Moral Philosophy. In *The Constitution of Agency: Essays on Practical Reason and Moral Psychology*, pp. 302–326. Oxford: Oxford University Press.
- Kripke, S. (1980). *Naming and Necessity*. Cambridge, MA: Harvard University Press.
- Lewis, D. (1983). New Work for the Theory of Universals. *Australasian Journal of Philosophy* 61: 343–377.
- Little, M. (1994). Moral Realism II: Non-Naturalism. *Philosophical Books* 35: 225–232.

- Mackie, J.L. (1977). *Ethics: Inventing Right and Wrong*. London: Penguin Books.
- Mill, J.S. (1871/1998). *Utilitarianism*, edited by Roger Crisp. Oxford: Oxford University Press.
- Miller, A. (2013). *Contemporary Metaethics: An Introduction*, 2nd edn. Cambridge: Polity.
- Moore, G.E. (1903/1993). *Principia Ethica*, 2nd edn. Cambridge: Cambridge University Press.
- Parfit, D. (2011). *On What Matters*, vols. i, ii. Oxford: Oxford University Press.
- Putnam, H. (1975). The Meaning of “Meaning.” In *Mind, Language, and Reality*, pp. 215–271. Cambridge: Cambridge University Press.
- Quine, W.V.O. (1969). Epistemology Naturalized. In *Ontological Relativity and Other Essays*, pp. 69–90. New York: Random House.
- Railton, P. (1986). Moral Realism. *Philosophical Review* 95: 163–207.
- Railton, P. (1989). Naturalism and Prescriptivity. *Social Philosophy and Policy* 7: 151–174.
- Ramsey, F. (1978). Theories. In *Foundations: Essays in Philosophy, Logic, Mathematics and Economics*, edited by D.H. Mellow, pp. 101–125. London: Routledge.
- Schroeder, M. (2007). *Slaves of Passions*. Oxford: Oxford University Press.
- Schroeder, M. (2008). *Being For: Evaluating the Semantic Program of Expressivism*. Oxford: Oxford University Press.
- Shafer-Landau, R. (2003). *Moral Realism: A Defence*. Oxford: Oxford University Press.
- Shafer-Landau, R. (2006). Ethics as Philosophy: A Defence of Ethical Non-Naturalism. In *Metaethics After Moore*, edited by Terry Horgan and Mark Timmons, pp. 209–232. Oxford: Oxford University Press.
- Sidgwick, H. (1907). *The Methods of Ethics*. New York: Macmillan.
- Smith, M. (1994). *The Moral Problem*. Oxford: Blackwell.
- Sturgeon, N. (1988). Moral Explanations. In *Essays on Moral Realism*, edited by Geoffrey Sayre-McCord, pp. 229–255. Ithaca: Cornell University Press.
- Suikkanen, J. (2010). Non-Naturalism: The Jackson Challenge. *Oxford Studies in Metaethics* 5: 87–110.
- Wedgwood, R. (2007). *The Nature of Normativity*. Oxford: Oxford University Press.