

Introduction

G. E. Moore (1873–1958) was a central figure in early twentieth-century philosophy. Along with Russell and Wittgenstein, he pioneered analytic philosophy, and his argumentative technique, his intellectual example, and his characteristic philosophical concerns shaped the way several generations of philosophers approached their discipline. As a result, Moore's influence is difficult to exaggerate. Even if few contemporary philosophers self-consciously adhere to any distinctively Moorean tenets or methods, his legacy is deeply and permanently embedded in English-language philosophy. This is particularly true in ethics, where no philosopher had a greater impact in the first half of the twentieth century than did Moore (Darwall, 1989: 366; Baldwin, 1993: xxxvii; Horgan and Timmons, 2006: 1).

His *Principia Ethica* of 1903, arguably “the first work in analytical ethical philosophy” (Darwall, 2006: 18), restructured the field, and until the publication of John Rawls's *A Theory of Justice* in 1971, no single work in ethics was to have repercussions as profound. In *Principia* Moore undertook to formulate the basic questions of ethics with precision, to clarify the differences among them, and to define the fundamental concepts involved in, and specify the procedures appropriate to, answering them. He paid particular attention to the concept of goodness, famously rejecting any naturalistic account of it in favor of the view that it represents a simple, indefinable, and unanalyzable property. Largely because of *Principia*, for a half century metaethical disputes over the nature of moral judgment, the meaning of moral language, and the possibility, if any, of justifying ethical propositions dominated moral philosophy.

Moore's objectives in *Principia*, however, were not exclusively or primarily metaethical. There and in his *Ethics* of 1912, he also wanted to address two substantive questions: What kinds of things are good in themselves? and What kinds of actions ought we to perform? Accordingly, this survey of Moore's ethical thought examines not only his metaethical nonnaturalism (Section 1) but also his value theory, including his doctrine of organic wholes, his repudiation of hedonism, and his distinctive account of the most important goods and evils (Sections 2 and 3), and his thinking about right and wrong – in particular, his critique of egoism and subjectivism and his elaboration of a nonhedonistic variant of utilitarianism and the implications it has for individual conduct (Sections 4 and 5). Because in recent decades metaethics has ceased to rule moral philosophy, the relevance and importance of these often-neglected aspects of Moore's ethical thought have only increased. In particular, Moore's development of a normative theory that shares the consequentialism of classical utilitarianism while abandoning its value theory, that is sensitive to the limits of

our knowledge, and that stresses the importance of adherence to rules while retaining an act-consequentialist criterion of right is more pertinent than ever.

1 The Meaning of "Good"

"It appears to me," Moore writes in the opening sentence of *Principia Ethica*, "that in Ethics, as in all other philosophical studies, the difficulties and disagreements of which its history are full, are mainly due to a very simple cause: namely, to the attempt to answer questions, without first discovering precisely *what* question it is you desire to answer" (33).¹ As a result, philosophers are "constantly endeavoring to prove that 'Yes' or 'No' will answer questions, to which *neither* answer is correct, owing to the fact that what they have before their minds is not one question, but several, to some of which the true answer is 'No,' to others 'Yes'" (33). In *Principia* Moore aspires to a higher standard of clarity and analytic rigor.

In particular, he wants to address two questions, both of which are central to moral philosophy but which, he believes, have been confused with one another or with other questions (33–4, 166). The first is, What ought to exist for its own sake – that is, what are the things that have intrinsic value or are good in themselves? Second is the question, What ought we to do – that is, what actions are we either permitted or required to perform? The first question is more basic, Moore believes, because we cannot know what good conduct is or what we ought to aim at without knowing what is good. "Nothing is more certain," he writes, "from an ethical point of view, than that some things are bad and others good; the object of Ethics is, indeed, in chief part, to give you general rules whereby you may avoid the one and secure the other" (94). Knowing what is good, however, and what is not requires understanding what is meant by "good." That is, what property or quality are we attributing to something when we say that it is good? True, we often successfully use a word or a concept, or correctly identify the things that possess a certain property, even if we cannot give a fully satisfactory definition or account of the concept or property in question. Moore acknowledges this, but he believes that if we don't know what "good" means, then "the rest of Ethics is as good as useless from the point of view of systematic knowledge." In particular, our "*most general* ethical judgments" are unlikely to be valid, and we won't understand what counts as evidence for any ethical judgment (57, 192–3).

¹ Parenthesized page numbers in the text are from *Principia Ethica* (Moore 1993b) unless preceded by an "E," in which case they refer to Moore's *Ethics* (Moore 2005a). Page references to other works by Moore are preceded by date and to the work of other authors by name and date.

Moore thinks, then, that “the most fundamental question in all Ethics” is “how ‘good’ is to be defined” (57, 192). In asking what is meant by “good,” however, Moore is not concerned with its proper usage as determined by custom (58). Nor is he calling for a dictionary definition or an account of the different ways the word “good” can be used, and indeed he acknowledges that it can have different meanings (E 82; 1962: 89–90). He is interested in the concept of goodness or the object, idea, or property that the adjective “good” stands for when used in ethical contexts – for example, when someone says that friendship is good or that pain is bad or asks whether pleasure is good in itself or has intrinsic value. What, then, is goodness? What does it mean for something to be good?

Moore’s answer is famous: “Good” is indefinable. “If I am asked ‘What is good?’ my answer is that good is good, and that is the end of the matter. Or if asked ‘How is good to be defined?’ my answer is that it cannot be defined, and that is all I have to say about it” (58). This response may seem disappointing, but Moore thought it had an important implication. If “good” cannot be defined – if it is an unanalyzable concept – then meaning alone cannot settle the truth or falsity of propositions about what is good or bad. Any such propositions are therefore synthetic, not analytic; that is, they are substantive assertions and never simply definitional or conceptual truths. This is important because, Moore thinks, philosophers frequently base their claims about what is good or bad on implicit or explicit appeals to definition. They confuse the question “What things are good in themselves?” with the question “What is meant by ‘good’?” (89–90, 109–10). But if good is indefinable, he writes, then nobody can foist on us propositions about what is good “on the pretence that this is ‘the very meaning of the word’” (58–9). We approach questions of what is good or bad with a more open mind, Moore believes, if we realize that this is not something that can be settled by definition.

Moore draws a parallel with “yellow.” Like it, “good” is a simple notion – too simple to be defined or analyzed. This does not imply that we don’t know what yellowness or goodness is or that we cannot recognize that something is “yellow” or “good.” To the contrary, we know perfectly well that the bananas in front of us are yellow or that kindness is good. But what we cannot do is define those terms or analyze or break them down into component parts. On Moore’s view, “definitions which describe the real nature of the object or notion denoted by a word . . . are only possible when the object or notion is something complex” (59). Like “yellow,” however, “good” does not refer to an object composed of parts that we can substitute in our minds when thinking of it. “The most important sense of ‘definition,’” Moore writes, “is that in which a definition states what are the parts which invariably compose a certain

whole; and in this sense ‘good’ has no definition because it is simple and has no parts” (61). Many other objects of thought are equally incapable of definition for the same reason. They “are the ultimate terms by reference to which whatever is capable of definition must be defined” (61). Although “good” is indefinable, nevertheless “we are all aware of a certain simple quality,” and it is this, and nothing else, that “we mainly mean by the term ‘good’” (90).

1.1 The Naturalistic Fallacy

Goodness is a property of certain natural objects or states of affairs, but it is not itself a natural property. It is not part of the natural world. It is not a constituent part of any physical object or any real state of affairs; that is to say, any physical object or state of affairs can be fully and completely described without mentioning its goodness or badness. When this point is understood, it seems less odd for Moore to contend, as he does, that the adjective “good” belongs to that “class of objects or properties of objects, which certainly do not exist in time, are not therefore parts of Nature, and which, in fact, do not *exist* at all” (161; cf. 93, 176). Various good things do exist, of course. However, only these things – the things that are good – and not goodness itself can exist in time, that is, have duration and begin or cease to exist. Only they can be objects of perception. In contending that the property designated by “good” does not exist in a material sense, Moore is not maintaining that it is imaginary. Goodness is like the number two. “Two and two *are* four. But that does not mean that either two or four exists” (162). There can be two real things – two hats, two cars, or two shoes – but “twoness” itself is not an object that exists.

Thus, when Moore denies “that ‘good’ *must* denote some *real* property of things,” he is not denying that it designates something (191). His point, rather, is that what it denotes is not a natural or empirical aspect of a thing. When we see that something is good, its goodness is “not a property which we can take up in our hands, or separate from it even by the most delicate scientific instruments” (175). As a nonnatural property, goodness lies beyond the province of psychology or the natural sciences; in ascribing goodness to an object, we do not describe the object at all (92; 1952: 591). It follows from the intrinsic nature of an object, but it is not one of its intrinsic properties; it does not make the object what it is (1993a: 296). Nor does “good” designate some supersensible property or refer to some metaphysical reality. “Any truth which asserts ‘This is good in itself’ is quite unique in kind . . . it cannot be reduced to any assertion about reality, and therefore must remain unaffected by any conclusions we reach about the nature of [physical or metaphysical] reality” (165; cf. 174).

Although “good” denotes a nonnatural property, and thus ethics is not a purely empirical enterprise, as was said above the things that are good do exist in the natural world. The adjective *good* must not be confused with the things to which the adjective applies, and those things are real existing objects or states of affairs. To say, for example, that “friendship is good” is not to define “good” as “friendship,” but to attribute a property, namely, goodness, to something that does or can exist, namely, friendship. Although we cannot define “good,” it does not follow from this “that *the* good, that which is good, is thus indefinable” (60). We can define the good by enumeration or extension, that is, by listing or describing those things that are good.

To attempt to define “good” and thus fail to recognize that it denotes a unique, indefinable, and unanalyzable quality, which is not equivalent to any other property, is to commit what Moore famously called the “naturalistic fallacy.” This fallacy “consists in identifying the simple notion which we mean by ‘good’ with some other notion” (109) – in particular, in contending “that good *means* nothing but some simple or complex notion, that can be defined in terms of natural qualities” (125). In trying to define what cannot be defined, the fallacy rests on a failure “to perceive that the notion of intrinsic value is simple and unique” (222; cf. 111). It is because goodness is frequently but erroneously identified with some natural or empirically determinable object or property like pleasure or that which is approved or desired that Moore calls the fallacy in question the naturalistic fallacy. But metaphysical theories of ethics commit the same fallacy when they identify good with some supersensible property or reality (91, 164–5). Goodness is distinct from any other property, whether empirical or metaphysical – that, Moore later stressed in an unpublished manuscript, was the point he was trying to drive home in *Principia* (13, 15).

There are, Moore contends, only two alternatives to the proposition that “good” denotes something simple and indefinable: Either it stands for “a complex, a given whole, about the correct analysis of which there may be disagreement; or else it means nothing at all, and there is no such subject as Ethics” (66). Let us begin with the second possibility. The passage in question is difficult to interpret, but what Moore seems to have in mind is not that “good” would be literally meaningless, but rather that if goodness were identical with some simple, natural property like pleasure, then “good” would lack any distinct or independent meaning. But “good” does have, he thinks, just such a meaning; it refers to a unique property or characteristic of things. If we ask whether something is good, we are not asking, for example, whether it is pleasant, and certainly if we ask whether pleasure is good, we are not wondering whether pleasure is pleasant. To the contrary, “good” has a distinct meaning for people even if they do not recognize the respect in which it is distinct: “Whenever [one]

thinks of ‘intrinsic value,’ or ‘intrinsic worth,’ or says that a thing ‘ought to exist,’ he has before his mind the unique object – the unique property of things – which I mean by ‘good.’ Everybody is constantly aware of this notion, although he may never become aware at all that it is different from other notions of which he is also aware” (68). Even if “good” can be used in different ways in different contexts (“she has a good credit score” or “chocolate sorbet tastes good”), it still has this core ethical meaning (1952: 554).

1.2 The Open-Question Argument

Turning to the other alternative – that “good” designates some complex property – Moore defends his contrary view that it is simple, indefinable, and unanalyzable by advancing what has come to be called the “open-question argument.” The argument is short and straightforward: “Whatever definition [of good] be offered, it may be always asked, with significance, of the complex so defined, whether it is itself good” (67). Although Moore is attacking definitions that identify good with some complex property, his argument applies equally well to definitions that equate it with some simple property. To understand the argument, suppose someone defines “good” as – to pick some possible examples – (1) pleasure, (2) that which promotes the survival of the species, (3) what God wants for us, or (4) what one would desire upon informed reflection. It is still perfectly coherent, Moore is arguing, for one to say, “I know that this is pleasurable (or what God wants for us or what an informed person would desire upon reflection), but is it good?” The very fact that a person can intelligibly ask whether a certain pleasure is good or whether what promotes the survival of the species is good – the very fact that these are open or meaningful questions – shows that we have a different notion in mind when we ask of something whether it is good than we do when we ask whether it is pleasurable or promotes the survival of the species or is what an informed person would desire.

By contrast, suppose someone said, “I know that Charles is a widower, but was he ever married?” That is not an open question. To the contrary, it shows only that the speaker has not grasped the meaning of the word “widower.” Of course, not every definition is as trivial as this. For example, many philosophers have held that “knowledge” means “justified true belief,” so that one knows X if and only if one believes X, X is indeed true, and one is justified in believing X. Although few epistemologists today accept this definition without qualification, assume, for the sake of discussion, that it is basically correct. If we have never thought about the meaning of “knowledge” before, this definition will come as new information to us and so it might seem to us an open question whether it is correct. But the more we think about it, the more convinced we may

become that “justified true belief” encapsulates what we mean by “knowledge.” If so, asking whether knowledge is justified true belief ceases to be an open question. When it comes to “good,” however, the situation is different. When we reflect on any proposed definition of it, we will see, as we mull it over, that it fails to capture what we mean when we call something “good.”

If we acknowledge that goodness is indefinable and, in particular, that it is distinct from any natural or metaphysical property, then, Moore believes, we will avoid the tangles that most ethical theories find themselves in. We will be less likely to err in our search for those things that are good because we will understand that this is not something that can be settled by definition. Moreover, if we try to rest our ethical principles on a definition of good, we cannot logically defend them. All we can do is arbitrarily assert (or deny) that “good” really means such-and-such. What is worse, we shall “be inclined either to misunderstand our opponent’s arguments or to cut them short with the reply, ‘This is not an open question: the very meaning of the word decides it; no one can think otherwise except through confusion’” (72).

Does the open-question objection, then, convincingly refute naturalism? A long-standing criticism of Moore’s argument is that even if it squelches this or that proposed naturalistic definition of good (as, say, pleasure, or what we would desire if fully informed), we cannot generalize from this and jump to the conclusion that any definition that someone might come up with will succumb to it. But Moore can be seen, not as trying to establish this conclusion by induction, but rather as laying down a challenge: Try to find a definition of good that gets around the open-question objection; you won’t succeed. As we examine various proposals, he thinks, we will come to realize that “good” escapes definition or analysis. Of course, it is possible that we might be surprised, as we were in the case of “knowledge,” by some new definition that we had never thought of before and that captures exactly what we have in mind when we call something good. But this seems unlikely. The concept of goodness has evaluative force; to describe a thing or state of affairs as good is to commend it, to attribute value to it, to say of it that it is better for it to exist than not to exist. And if something like this is correct, then no descriptive definition can fully capture the meaning of good.

Most philosophers today agree with Moore that “good” is not synonymous with or conceptually equivalent to any nonmoral concept. Still, it might possibly be the case that the property of goodness is, in fact, identical to some natural property X even if “good” does not mean “X.” If so, then some kind of nondefinitional or nonanalytic naturalism would be true. Some contemporary philosophers defend this position, which seems safe from the open-question argument, by drawing analogies with scientific knowledge. Consider salt, for

example. Salt, we now know, is NaCl, but this was a discovery about salt. It was not a definitional truth, which we learned by reflecting on our concepts, because “salt” and “NaCl” have different meanings. Likewise, these contemporary naturalists urge, goodness is (or might be) identical to or reducible to some natural property or properties even if this is not an analytic or conceptual truth.

The tenability of this proposal and, more specifically, the analogy with science are matters of debate. For one thing, our prescientific concept of salt left open for empirical investigation the true nature of the substance that has the properties we associate with salt. But it is doubtful that our concept of goodness is similarly incomplete, waiting to be filled out by further research. Moore himself did not explicitly consider nonanalytic naturalism, but its contemporary critics reject it on broadly Moorean grounds, arguing that no account of goodness or any other ethical property that reduces it to or identifies it with some natural property, whether by definition or otherwise, can capture its normative (or commendatory or action-guiding) force. Even if everything that is X is good, and the only things that are good are X, still the property of being good would seem to be distinct from the property of being X. In other words, even if it were true that X and only X makes things good, their coextensiveness would not entail that X-ness and goodness are the very same thing (Parfit, 2011–7, III: 71–2, 75–83).

Moore believed that most earlier writers in ethics were guilty of the naturalistic fallacy – of identifying goodness with some natural or metaphysical property – and he devotes nearly three chapters of *Principia* to showing how thinkers as diverse as Aristotle, Spinoza, Leibniz, Kant, Bentham, Mill, Green, and Spencer rested their ethical theories on some implicit or explicit claim about what good means. Henry Sidgwick, the great utilitarian thinker and one of Moore’s teachers, is the only philosopher acknowledged by Moore to have understood that good is unanalyzable. But although the writers Moore criticizes do link the good to different natural or supersensible properties, it is far from clear – perhaps even to these writers themselves – whether they are doing so on the basis of a definition of “good.” As Moore himself stresses, one does not commit the naturalistic fallacy merely by identifying the kinds of thing that are good or claiming that X is the only good. However, a substantive assertion about what is good requires some sort of support or argument to back it up, and Moore thought, correctly or incorrectly, that these writers were not offering arguments but passing off their views about the good as definitional truths.

Whatever the merits of Moore’s critique of these philosophers, *Principia Ethica* profoundly influenced the course of metaethics in the twentieth century. Before too long it was recognized as a classic, and many philosophers found Moore’s bracing critique of naturalism fully convincing. However, his doctrine

that ethical knowledge involves the apprehension of a *sui generis*, nonnatural property (namely, goodness) never found as much favor. Many philosophers were uncomfortable with the idea that we are able, somehow, to intuit, recognize, or apprehend certain nonnatural ethical properties or to grasp in a nonempirical way the truth of propositions referring to those properties. This can be seen as an epistemological worry. They were also puzzled about the ontological status of nonnatural properties and how, exactly, they are related to the natural (or descriptive or nonmoral) properties on which they supervene. Moore thought that if two things have the same empirical properties, then they will be good or bad to the same degree. Their goodness can differ only if their empirical or descriptive properties differ. But how, then, do those natural properties give rise to nonnatural properties? This can be seen as a metaphysical worry.

Worries like these may have reflected inaccurate understandings of Moore, who did not believe that ethical knowledge was the product of a distinct kind of cognition (36) and avoided attributing some special reality to nonnatural properties (162–3). Nevertheless, they led some philosophers, not to try to breathe life into naturalism, but rather to adopt *noncognitivism*, generally understood as the position that distinctively ethical discourse is not cognitive at all – that is, that ethical statements are neither true nor false. Rather than being propositions with truth value (for example, the proposition that X has the property of goodness), ethical statements are an expression of attitude or emotion or some kind of prescription or endorsement. To say “X is good” is not to say something that is either true or false but, rather, to recommend, endorse, prescribe, or express some other positive attitude toward X. Noncognitivists accept Moore’s open-question argument, but reject his cognitivism, his belief that there are true ethical propositions and that we can know some of them.

Noncognitivism dominated Anglo-American philosophy in the middle of the twentieth century and still has influential and philosophically sophisticated advocates. But these days many philosophers reject it, largely because moral discourse and the structure of moral language seem to presuppose that ethical statements (e.g., “torture is wrong”) make claims about the way things are, claims that can be true or false and are not just prescriptions (“don’t torture”) or expressions of attitude (“I’m against torture”).² Some of these philosophers favor some kind of nonanalytic naturalism and seek to show that moral properties and moral facts, even if not definitionally or conceptually equivalent to

² To be sure, noncognitivists try to explain why we are justified in acting as if moral judgments had truth value even though, on their view, they don’t. So-called moral error theorists (Mackie 1977) take a different tack. They grant that moral judgments have truth value, but contend that all such judgments are false.

natural properties and facts, are nevertheless identical to them. Other contemporary cognitivists are, like Moore, nonnaturalists. They believe that we have moral knowledge but that moral properties and facts cannot be reduced to, or identified with, natural or nonethical properties or facts. Thus, metaethics since *Principia* can be seen as an increasingly intricate and sophisticated three-way debate among proponents of naturalism, of nonnaturalism, and of noncognitivism – a debate set off by Moore's open-question argument.

2 Intrinsic Goodness

Throughout his writings, Moore emphasizes that judgments about goodness divide into two types: judgments about whether something is good in itself and judgments about whether it is good as a means – that is, as a cause or necessary condition of something else that is good in itself. “The nature of these two species of universal ethical judgments is extremely different,” he states in *Principia*, “and a great part of the difficulties, which are met with in ordinary ethical speculation, are due to the failure to distinguish them clearly” (73; cf. 75). In particular, we need to avoid the error of supposing that something that is a means to good and, indeed, which may even seem necessary here and now for the existence of anything good, is therefore good in itself (236). Appreciating the difference between these two types of judgments, Moore contends, is absolutely crucial if one is to analyze and answer correctly potentially ambiguous questions like “What should we aim at?” or “Is it right to act this way?” To answer such questions, “we must know *both* what degree of intrinsic value different things have, *and* how these different things may be obtained” (77).

Judgments about goodness as a means are essentially causal and predictive. They assert that some action, thing, or state of affairs will have certain effects. However, because circumstances vary, it is virtually impossible that a given thing or action will always produce the same result. Indeed, “a thing which has good effects under some circumstances may have bad effects under others” (78–9). The most we can hope to know is that a certain result “generally follows” this kind of thing (74). And even this generalization will hold only if circumstances are generally the same, and although this may be so for a particular age or at a certain stage of society, what is generally true at one period may be generally false at another. Moreover, to judge that something is generally a means to good we need to know not only that it usually does some good but also that the balance of good will generally be greater than if one had done something else instead (74).

By contrast, judgments that certain kinds of things or states of affairs are good in themselves – that they are intrinsically valuable – are not causal or empirical