

DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD
DASH.HARVARD.EDU

HARVARD LIBRARY


Two Distinctions in Goodness

Citation
Korsgaard, Christine. 1983. Two distinctions in goodness. Philosophical Review 92, no. 2: 169-195.

Published version
https://doi.org/10.2307/2184924

Link
http://nrs.harvard.edu/urn-3:HUL.InstRepos:3164346

Terms of use
This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material (LAA), as set forth at
https://harvardwiki.atlassian.net/wiki/external/NGY5NDE4ZjgzNTc5NDQzMGIzZWZhMGFlOWI2M2EwYTg

Accessibility
https://accessibility.huit.harvard.edu/digital-accessibility-policy

Share Your Story
The Harvard community has made this article openly available.
Please share how this access benefits you. Submit a story

The Philosophical Review, XCII, No. 2 (April 1983)

TWO DISTINCTIONS IN GOODNESS

Christine M. Korsgaard

In this paper I describe two distinctions in goodness which are often conflated, and try to show the importance of keeping them separate. The two distinctions in question are: the distinction between intrinsic and extrinsic goodness, and the distinction between ends or final goods, and means or instrumental goods.

It will help to begin by delineating the kind of value and the kind of judgment of value with which I am primarily concerned here. I take it that there are three primary categories of value with which the moral philosopher is concerned: namely, the rightness or justice of actions, policies, and institutions; the goodness of objects, purposes, lives, etc.; and the moral worth or moral goodness of characters, dispositions, or actions. My concern here is not with what constitutes moral worth or moral goodness but with the second category—with goodness as a feature of ordinary ends and purposes, states of affairs, objects, activities, and other things—that is, with the kind of goodness that marks a thing out as worthy of choice.

Within this category, we can distinguish, admittedly with some artificiality, three kinds of judgments of goodness that we make. We judge something to be good of its kind when we judge it to have the virtues appropriate to that kind. We may also judge something to be a good kind of thing, as when we say of friendship or books or health that they are good. And we also sometimes judge particular things to be good absolutely, meaning that here and now the world is a better place because of this thing. I am mostly concerned with this third sort of judgment in this paper, though part of what is in question is its relation to the other two.

II

It is rather standard fare in philosophy to distinguish two kinds of this value of goodness, often called "intrinsic" and "instrumen-

tal."[1] Objects, activities, or whatever, have an instrumental value if they are valued for the sake of something else—tools, money, and chores would be standard examples. A common explanation of the supposedly contrasting kind, intrinsic goodness, is to say that a thing is intrinsically good if it is valued for its own sake, that being the obvious alternative to a thing's being valued for the sake of something else. This is not, however, what the words "intrinsic value" mean. To say that something is intrinsically good is not by definition to say that it is valued for its own sake: it is to say that it has its goodness in itself. It refers, one might say, to the location or source of the goodness rather than the way we value the thing. The contrast between instrumental and intrinsic value is therefore misleading, a false contrast. The natural contrast to intrinsic goodness—the value a thing has "in itself"—is *extrinsic* goodness—the value a thing gets from some other source. The natural contrast to a thing that is valued instrumentally or as a means is a thing that is valued for its own sake or as an end. There are, therefore, two distinctions in goodness. One is the distinction between things valued for their own sakes and things valued for the sake of something else—between ends and means, or final and instrumental goods. The other is the distinction between things which have their value in themselves and things which derive their value from some other source: intrinsically good things versus extrinsically good things. Intrinsic and instrumental good should not be treated as correlatives, because they belong to two different distinctions.

If intrinsic is taken to be the opposite of instrumental, then it is under the influence of a theory: a theory according to which the two distinctions in goodness are the same, or amount to the same thing. According to such a theory, final goods or things valuable as ends will be the same as intrinsic goods, and instrumental goods or things valuable as means will be the same as extrinsic goods. It is worth considering what such a theory might be like.

[1]Intrinsic is often directly opposed to instrumental; equally commonly, "extrinsic" is opposed to intrinsic but then "consequential" or "instrumental" is offered as a definition or explanation of that term. Or, in some of the literature, "intrinsic" is taken to be a particular theory about how ends are valued, and accepted or dismissed as such. All of these usages more or less imply the equivalence of the two distinctions; none leaves room for the Kantian theory described in this paper.

The first part of the equivalence—that ends and intrinsic goods are the same—might be held in two very different ways: (1) The claim might be that anything we value for its own sake is thereby "intrinsically good"; that is, that this is all that can be meant by "intrinsically good." This amounts to a *reduction* of the intrinsic/extrinsic distinction to the end/means distinction; the significance of the former distinction drops out. This option, which in effect replaces the intrinsic/extrinsic distinction with the end/means distinction, is sometimes taken to render the conception of good "subjective," both in the sense of 'relative to the person' and of 'varying among individuals'. The thought goes this way: good things (on this account) are just those valued for their own sakes, but different people value different things for their own sakes. (2) The second way one might equate ends and intrinsically good things is by claiming that those things which have intrinsic value are or ought to be treated as ends. In this case we have a significant, and rather metaphysical, claim about ethics and moral psychology: namely, that choice is or ought to be a response to an attribute that we perceive in things—the attribute of intrinsic goodness. Equating the two distinctions in goodness thus leads naturally to the idea that there are two alternative theories about final goods—either that "good" is subjective or that good things are the possessors of some particular attribute. Objectivity, in other words, is thought to amount to the possession of an attribute. I think that many people do have a tendency to think that these alternatives are exhaustive, and one thing I want to show is that if the two distinctions in goodness are kept separate, this need not be so.

The other side of the theory that equates the two distinctions is the equation of extrinsic with instrumental goods, or means. The consequences of this equation are serious. Since intrinsically good things (at least when "intrinsic" retains its significance) are thought to have their value *in* themselves, they are thought to have their goodness in any and all circumstances—to carry it with them, so to speak. If you find that a certain kind of thing is not good in any and all circumstances, that it is good in some cases and not others, its goodness is extrinsic—it is derived from or dependent upon the circumstances. If extrinsic value and instrumental value are equated, you are then forced to say of all such things that they are means or instruments. This way of thinking is part of what is be-

hind the tendency to conclude that the final good must be pleasure or some sort of experience. The argument proceeds as follows: take an activity that we would naturally say is valuable for its own sake,—say, looking at a beautiful sunset. Now the question is raised: would you think this activity was a good one even if the person engaged in it found it tedious or painful? If you say "no" then you have admitted that the goodness of this activity is not intrinsic; that it depends, in some way, on the pleasantness of it. But if all extrinsic value is instrumental value, then the only option is that the activity is a *means* to pleasure. Now if the two distinctions are not equated, there is room for some other sorts of accounts of extrinsic value, and one may not be forced to this conclusion.

Because of these consequences this side of the equation has been more widely attacked than the other. It has been argued that instrumentality is not the only sort of extrinsic value, on the grounds that there are other sorts of contributions things can make to intrinsically good ends. So, for instance, it is common to identify a "part" of an intrinsically valuable "whole" as having "contributive" value. Another sort of value, suggested by C. I. Lewis, is called "inherent" value.[2] This is supposed to be the value that characterizes the object of an intrinsically good experience. A painting, for example, might have inherent value. The identification of these different kinds of extrinsic value serves as a reminder that things can bear other relations to good ends besides being their causes or tools for their production. Contributive value and inherent value, however, both share with instrumental value the fact of deriving their goodness from the contribution they make to the existence of a supposedly intrinsically good end.

Separating the two distinctions in goodness, however, opens up another possibility: that of something which is extrinsically good yet valued as an end. An example of this would be something that was good as an end because of the interest that someone took in it, or the desire that someone had for it, for its own sake. This is the case that I am going to be discussing in the rest of this paper. In particular, I am going to compare the very opposite treatments of this issue that appear in, on the one hand, the work of Moore and Ross at the beginning of our century and, on the other hand, Kant.

[2]C. I. Lewis, *An Analysis of Knowledge and Valuation* (Open Court, 1946).

These philosophers all separated the two distinctions, but they applied them to this case in opposite ways. Moore and Ross came to the conclusion that the goodness of ends is intrinsic and must be independent of the interest that people take in them or the desires that people have for them. You might value something as an end because of its intrinsic goodness or in response to its intrinsic goodness, but a thing's possession of intrinsic goodness is quite independent of whether anyone cares about it or not. Kant's theory, on the other hand, both allows for and depends upon the idea of extrinsically valuable ends whose value comes from the interest that people take in them.

The fact that philosophers nowadays often oppose intrinsic to instrumental value and equate intrinsic value with the value of ends may just be taken to be sloppiness, of course. But it may also mean that these philosophers are working with some theory of the sort I have described—a theory of the equivalence of the two distinctions. As the Kantian option shows, such a theory is a substantive philosophical position and restricts the possibilities open to us in serious ways. It should not, in any case, be taken for granted.

III

In the early years of this century there was much discussion of the question whether or not a good thing has its value as a result of something like the interest taken in it or the desire someone has for it. Influential philosophers such as G. E. Moore, W. D. Ross, R. B. Perry, and others discussed this question at length. Probably the interest in the issue was aroused by a common utilitarian argument that pleasure is the only thing that is good in itself because it is the only thing that we can desire for its own sake. It quickly became, and still is, a commonplace in discussions of utilitarianism to argue that pleasure is not, after all, the only thing that we desire for its own sake. But that leaves open the further question whether the things we desire for their own sakes, whatever they might be, are therefore good in themselves, or intrinsically good. Moore, and following him, Ross, argued vigorously that this could not be so. Goodness, they said, had nothing to do with mental attitudes taken towards things at all—even though it turned out that, as a matter of

fact, goodness is a property of mental attitudes or a property of states of affairs that always include mental states or attitudes.

The idea of intrinsic value is central to Moore's theory. He believed that right actions are those that maximize intrinsic goods. Emphatically opposed to hedonism, he took the class of intrinsic goods to consist of such things as the appreciation of beauty, friendship, and love. In his attempt to account for the goodness of these things, he came back to the question of the nature of intrinsic value over and over again.[3]

In his paper "The Conception of Intrinsic Value," Moore argues that people who object to the idea that goodness is subjective are really worried about something quite different: the idea that goodness is nonintrinsic. This is shown, according to Moore, by the fact that there are theories which would render goodness objective to which the same people would still be opposed, and for the same reason. Moore gives as his example the theory that "better" means "better fitted to survive":[4] people who object to subjectivity, he says, would also object to this, although it renders "good" objective. So the problem with a subjectivist theory of the good is not merely the lack of objectivity, but something else. According to Moore, it is that it excludes the possibility that things are intrinsically valuable. Moore defines intrinsic value as follows:

> To say that a kind of value is "intrinsic" means merely that the question whether a thing possesses it, and in what degree it possesses it, depends solely on the intrinsic nature of the thing in question.[5]

Moore's definition of the intrinsic nature of a thing is rather complicated: he says that two things have a different intrinsic nature if they are not exactly alike; that the difference need not be a dif-

[3]Moore's views on intrinsic value are mostly presupposed in *Principia Ethica* (Cambridge, 1903); but they are addressed explicitly in *Ethics* (Oxford, 1912); "The Conception of Intrinsic Value" written for *Philosophical Studies* (Routledge and Kegan Paul, Ltd., 1922); and a symposium reply entitled "Is Goodness a Quality?" published in the Aristotelian Society Supplement in 1932 and reprinted in *Philosophical Papers* (1959). In the last, Moore tends to give way to a view that his earlier accounts avoid—namely, that only experiences can be intrinsically good. For that reason and because of its polemical nature I have not used it in this paper.

[4]"The Conception of Intrinsic Value" p. 256.

[5]*Ibid.*, p. 260.

ference in qualities, since it may be in the degree of a quality or in the quality of a constituent; and that two numerically different things have the same intrinsic nature if they are exactly alike.[6] In general, the intrinsic nature of a thing seems to consist of its nonrelational properties, for Moore insists that a thing would have the same intrinsic nature if transferred to another world or placed in a different set-up of causal laws.[7] This is what Moore supposes we *want* from the conception of intrinsic goodness, as his analysis of the trouble with the evolutionary account of goodness shows. He says that the difficulty is that the types better fitted to survive under our laws of nature would not be the same as the types better fitted to survive under other circumstances and with different laws of nature. "Good" therefore would not be dependent only on a thing's intrinsic nature but would be a property that is relative to the circumstances, even though in this case it would be objective. But the problem with subjectivism is the same: it makes "good" relative to the circumstances.

Intrinsic goodness is not an element in the thing's intrinsic nature, for to say that would be to commit the naturalistic fallacy. The elements in its intrinsic nature are natural properties and cannot be identified with the good. But it is dependent only on the thing's intrinsic nature and is just as constant: so long as the thing remains what it is, it has the same value: and the value is the same, of course, for everyone and so also objective. Since it is no part of a thing's intrinsic nature whether anybody likes it or not, intrinsic value is quite independent of people's desires and interests. To put it another way: the attribute of "being desired by somebody" is relational, and as such it obviously varies with the circumstances in which the thing is found.

In *Ethics,* Moore's definition is a little different. We judge a state of things to have intrinsic value when we judge that it would be a good thing for that state of things to exist, even if nothing else were to exist besides. Here again, the emphasis is on the thing's goodness being nonrelational in a certain way. This view of intrinsic goodness is behind Moore's method of determining which things have intrinsic goodness in *Principia Ethica:* the "method of isolation." In

[6]*Ibid.,* pp. 260–65.
[7]*Ibid.,* p. 256.

order to arrive at a correct decision on the question which things have intrinsic value, Moore says that we must consider whether a thing is such that, if it existed by itself, in absolute isolation, we should judge its existence to be good.[8] In *Ethics,* Moore says:

> We *can* consider with regard to any particular state of things whether it would be worth while that it should exist, even if there were absolutely nothing else in the Universe besides . . . we *can* consider whether the existence of such a Universe would have been better than nothing, or whether it would have been just as good that nothing at all should ever have existed.[9]

These definitions, along with the method of isolation they suggest, seem to Moore to exclude easily any connection between intrinsic value and what people desire for its own sake, for, he tells us, it is obviously possible to desire something for its own sake, or believe that someone else does, and yet not regard the thing as the sort of thing that would be good if it existed in isolation. Indeed you might regard it as a bad thing, worse than nothing, for it to exist quite alone. Moore concludes:

> And if this is so, then it shows conclusively that to judge that a thing is intrinsically good is not the same thing as to judge that some man is pleased with it or desires it for its own sake.[10]

Moore, it should be noted, does not usually use the terminology of "relational" vs. "nonrelational" attributes in his discussions of intrinsic value, but these are the terms in which Ross and Perry, following Moore, take up the discussion. Ross, who is on Moore's side,[11] says that there are two kinds of theories of value. One kind treats value as an attribute, and the other treats it as a relation, usually to a state of mind such as interest or desire. If it is a relation,

---

[8]*Principia Ethica,* p. 187.

[9]*Ethics,* p. 68.

[10]*Ibid.,* p. 69.

[11]In *The Right and The Good* (Oxford, 1930) Chapter IV, Ross argues explicitly in favor of Moore and against Ralph Barton Perry, who, in his *General Theory of Value* (Harvard, 1926), argues that value is relative to interest. The Kantian view defended in this paper is classified by Perry as one in which value is "the object of a qualified interest" and opposed by him in favor of the view that value is "the object of any interest."

Ross complains, then nothing can be intrinsically good, since intrinsically good *means* "good even if nothing else exists." But, he says,

> . . . in that case value would seem always to be borrowed, and never owned; value would shine by a reflected glory having no original source.[12]

Ross, like Moore, finds it virtually self-evident that "intrinsically valuable" is not the same as "desired as an end." He insists that:

> It is surely clear that when we call something good we are thinking of it as possessing in itself a certain attribute and are not thinking of it as necessarily having an interest taken in it.[13]

The terms in which this discussion proceeded suggested that the question was whether final goods, whatever we ought to pursue, are intrinsically good and objective, the possessors of a property; or good because they are desired and therefore subjective, or at any rate "relational" and therefore unfixed. These are terms that those who followed Moore and Ross inherited.

IV

Kant, I am going to claim, was aware of and made use of the two distinctions in goodness, with results that were quite different from those arrived at by Moore and Ross.[14] In order to see this, we must

---

[12]*The Right and the Good,* p. 75.

[13]*Ibid.,* p. 81.

[14]I am not the first to set up Kant's view in opposition to Moore's. The same is done by H. J. Paton in "The Alleged Independence of Goodness" written for the Library of Living Philosophers Volume on Moore (Northwestern: Volume 4, 1942). Paton, however, is not concerned with the two distinctions, and he focuses on the goodness of actions, which he claims is relative to the circumstances in which they are performed. Moore's rather impatient response is to incorporate the choice into the action and consideration of the circumstances into the choice: thus under different circumstances you have different actions. Moore's reply may be fair in the case discussed, but it is an instance of a general strategy which I discuss in the paper: when someone brings forward an example of a good thing whose goodness seems relative to the circumstances, Moore and Ross incorporate the circumstances into the thing to maintain the nonrelational character of the goodness.

begin by looking at Kant's own distinction between unconditioned and conditioned value. The *Foundations of the Metaphysics of Morals* opens with the famous claim:

> Nothing in the world—indeed nothing even beyond the world—can possibly be conceived which could be called good without qualification except a *good will* (G 9/392–393).[15]

As Kant presents the argument that follows, it becomes clear that what he means is that the good will is the only unconditionally good thing and "the supreme condition to which the private purposes of men must for the most part defer" (G 12/396). He says:

> This will must indeed not be the sole and complete good but the highest good and the condition of all others, even of the desire for happiness (G 12/396).

Happiness, by contrast to the good will, is referred to as a "conditional purpose" (G 12/396).

The fact that happiness is identified as a conditional purpose shows that the unconditioned/conditioned distinction is not the same as the end/means distinction, since happiness is certainly desired as an end. For Kant, the end/means distinction can be said to be a distinction in the *way* we value things. By contrast, the unconditioned/conditioned distinction is a distinction not in the way we value things but in the circumstances (conditions) in which they are objectively good. A thing is unconditionally good if it is good under any and all conditions, if it is good no matter what the context. In order to be unconditionally good, a thing must obviously carry its own value with it—have its goodness in itself (be an end in itself). Kant's notion of unconditional value therefore corresponds to the notion of intrinsic goodness as nonrelational that I have been discussing. The early passages of the *Foundations* emphasize the inde-

[15]References to Kant's works are given in the text as shown. "G" stands for the *Foundations of the Metaphysics of Morals;* the first page number is that of the translation by Lewis White Beck (Library of Liberal Arts, 1959); the second is the Prussian Academy edition page number. Other titles are given in full. The translations used are: Lewis White Beck, *Critique of Practical Reason* (Library of Liberal Arts, 1956) and Mary J. Gregor, *The Doctrine of Virtue* (Harper Torchbooks, 1964).

pendence of the value of the good will from all surrounding circumstances as well as from its results. It is good in the world or even beyond it (G 9/393); it is not good because of what it effects or accomplishes; it sparkles like a jewel in its own right, as something that has its full worth in itself. Later in the *Foundations,* Kant uses the phrase "*inneren Wert,*" inner worth, to describe the special dignity of a morally good rational being as compared to the "*relativen Wert,*" relative worth, of anything else (G 53/435). But whereas Moore assigned intrinsic goodness to a range of things—to aesthetic appreciation, to friendship, and in general to the things that he thought we ought to pursue as ends—Kant assigns it only to the one thing, the good will.

If unconditional value is intrinsic value, conditional value is extrinsic value. Now a thing is conditionally valuable if it is good only when certain conditions are met; if it is good sometimes and not others. Thus, to elaborate on Kant's own examples, "the coolness of a villain makes him not only far more dangerous but also more directly abominable in our eyes than he would have seemed without it" (G 10/344), while coolness in a fireman or a surgeon is usually an excellent thing. Power, riches, and health are good or not depending upon what use is made of them. To say that a thing is conditionally valuable is to say that it is good when and only when the conditions of its goodness are met. We can say that a thing is good objectively (this is my terminology) either if it is unconditionally good or if it is a thing of conditional value and the conditions of its goodness are met. Here it is important to notice that "good objectively" is a judgment applying to real particulars: this woman's knowledge, this man's happiness, and so on. To say of a thing that it is good objectively is not to say that it is the type of thing that is usually good (a good kind of thing like knowledge or happiness) but that it contributes to the actual goodness of the world: here and now the world is a better place for *this.* We would not say that about the coolness of the villain or the happiness of the evil person: hence coolness and happiness are objectively good only when certain further conditions are met. Further, we might, under unusual conditions, attribute objective goodness to something that under more usual conditions is nearly always bad, as when a kind of occurrence normally unfortunate coincidentally contributes to

someone's happiness.[16] When Kant says that the only thing good without qualification is a good will, he means that the good will is the one thing or kind of thing for which the world is always a better place, no matter "what it effects or accomplishes" (G 10/394).

The two distinctions interact in the following ways. When a thing is valued as a means or instrumentally (or is the sort of thing valued as a means) it will always be a conditionally or extrinsically valuable thing, and the goodness of the end to which it is a means will be a condition of its goodness. Instruments therefore can only be conditionally valuable. If the conditions of their goodness are met, however, they can be good objectively. The more important point is about things valued as ends. These are also conditionally or extrinsically good. In particular, happiness, under which Kant thinks all our other private purposes are subsumed, is only conditionally good, for:

> It need hardly be mentioned that the sight of a being adorned with no feature of a pure and good will, yet enjoying uninterrupted prosperity, can never give pleasure to a rational impartial observer. Thus the good will seems to constitute the indispensable condition even of worthiness to be happy (G 9/393).

But although happiness is conditionally valuable, it is, when the condition is met, objectively good.

In order to see this, it will help to keep in mind Kant's other uses of the unconditioned/conditioned distinction. If anything is conditioned in any way, reason seeks its condition, continually seeking the conditions of each condition until it reaches something unconditional. It is this characteristic activity of reason that generates the antinomies of theoretical speculative reason described in the *Critique of Pure Reason*. The usual example is causal explanation—if we explain a thing in terms of its cause, we then go on to explain the cause itself in terms of its cause, and this process continues. Reason does not want to rest until it reaches something that needs no explanation (although this turns out not to be available): say, something that is a first cause or its own cause. A causal explanation truly satisfying to reason would go all the way back to this evident first cause, thus *fully* explaining why the thing to be explained must

[16] I am indebted to the Editors of the *Philosophical Review* for this point.

be so. These are familiar sorts of moves in philosophy, so there is no need to belabor the point. To apply it here, it is only necessary to point out that just as to explain a thing fully we would have to find its unconditioned first cause, so to *justify* a thing fully (where justify is "show that it is objectively good") we would have to show that all the conditions of its goodness were met, regressing on the conditions until we came to what is unconditioned. Since the good will is the only unconditionally good thing, this means that it must be the source and condition of all the goodness in the world; goodness, as it were, flows into the world from the good will, and there would be none without it. If a person has a good will, then that person's happiness (to the extent of his or her virtue) is good. This is why the highest good, the whole object of practical reason, is virtue and happiness in proportion to virtue: together these comprise all ends that are objectively good—the unconditional good and the private ends that are rendered good by its presence. (*Critique of Practical Reason,* 114–115/110). So also the Kingdom of Ends, defined as "a whole of rational beings as ends in themselves as well as of the particular ends which each may set for himself" (G 51/433), is a kingdom in which the objectively good is fully realized.

On the Kantian conception of goodness, then, an end is objectively good either if it is unconditionally (intrinsically) good *or* if it is conditionally good and the relevant conditions, whatever they are, are met. This conception of the good is used in his argument for one of the formulas of the categorical imperative, the Formula of Humanity as an End in Itself.[17] It is this argument that establishes the role of the good will in conferring value upon the ends of the person who has it.

The argument shows how Kant's idea of justification works. It can be read as a kind of regress upon the conditions, starting from an important assumption. The assumption is that when a rational being makes a choice or undertakes an action, he or she supposes the object to be good, and its pursuit to be justified. At least, if there is a categorical imperative there must be objectively good ends, for then there are necessary actions and so necessary ends (G 45–46/427–428 and *Doctrine of Virtue* 43–44/384–385). In order

[17]A much fuller treatment of the ideas of this section is in my paper "Kant's Formula of Humanity," forthcoming in *Kant-Studien.*

for there to be any objectively good ends, however, there must be something that is unconditionally good and so can serve as a sufficient condition of their goodness. Kant considers what this might be: it cannot be an object of inclination, for those have only a conditional worth, "for if the inclinations and the needs founded on them did not exist, their object would be without worth" (G 46/428). It cannot be the inclinations themselves because a rational being would rather be free from them. Nor can it be external things, which serve only as means. So, Kant asserts, the unconditionally valuable thing must be "humanity" or "rational nature," which he defines as "the power set to an end" (G 56/437 and DV 51/392). Kant explains that regarding your existence as a rational being as an end in itself is a "subjective principle of human action." By this I understand him to mean that we must regard ourselves as capable of conferring value upon the objects of our choice, the ends that we set, because we must regard our ends as good. But since "every other rational being thinks of his existence by the same rational ground which holds also for myself" (G 47/429), we must regard others as capable of conferring value by reason of their rational choices and so also as ends in themselves. Treating another as an end in itself thus involves making that person's ends as far as possible your own (G 49/430). The ends that are chosen by any rational being, possessed of the humanity or rational nature that is fully realized in a good will, take on the status of objective goods. They are not intrinsically valuable, but they are objectively valuable in the sense that every rational being has a reason to promote or realize them. For this reason it is our duty to promote the happiness of others—the ends that they choose—and, in general, to make the highest good our end.

It is worth emphasizing that the relation of intrinsic to extrinsic value in this case—the case of extrinsically valuable ends—is entirely different from that in the cases of extrinsic value mentioned earlier. Instrumental value, contributive value, and Lewis's inherent value were all forms of extrinsic value that derived from the production of a supposedly intrinsically good end. The extrinsic value of an objectively good end—of something that forms part of the happiness of a good person—comes not from some further thing that that end promotes but from its status as the object of a rational and fully justified choice. Value in this case does not travel

from an end to a means but from a fully rational choice to its object. Value is, as I have put it, "conferred" by choice. This formulation may seem paradoxical. A natural objection will be that the goodness of the chosen object is precisely what makes the choice rational, so that the choice cannot itself be what makes the object good. I will have more to say about this objection in the next section (see pp. 184–190). The point I want to emphasize here is that the Kantian approach frees us from assessing the rationality of a choice by means of the apparently ontological task of assessing the thing chosen: we do not need to identify especially rational ends. Instead, it is the reasoning that goes into the choice itself—the procedures of full justification—that determines the rationality of the choice and so certifies the goodness of the object. Thus the goodness of rationally chosen ends is a matter of the demands of practical reason rather than a matter of ontology.[18] It is notable that on Kant's theory the goodness of means is handled the same way: it is not because of the ontological property of being productive of an intrinsically good end that means are good but rather because of the law of practical reason that "whoever wills the end, so far as reason has decisive influence on his action, wills also the indispensably necessary means to it that lie in his power" (G 34/417). Similarly, the argument for the objective goodness of the object of a rational choice is not an ontological one; rather, it is based on Kant's theory of rational action. If we regard our actions as rational, we must regard our ends as good; if so, we accord to ourselves a power of conferring goodness on the objects of our choice, and we must accord the same power—and so the same intrinsic worth—to others.

It will be helpful to pause for a moment to match up Kant's view and the Kantian terms to what has gone before. On Kant's view there is only one thing that has what he calls unconditional value and what Moore calls intrinsic value, and that is the power of rational choice (when the choices are made in a fully rational way,

[18]Insofar as Moore's point in identifying the naturalistic fallacy is to deny the identity of goodness with any particular natural property and so to insist on the autonomy of ethical discourse, Kant could agree. But whereas Moore concludes that goodness must therefore be a nonnatural property, Kant understands it to be a practical, rather than a theoretical, characterization.

which is what characterizes the good will). The value of everything else whatever is extrinsic or conditional. Yet when a thing is conditionally valuable and the relevant conditions are met, the thing has objective value. Things that are valued for their own sakes or as ends have this status. Their value is conditional but can be objective, given the real circumstances of the case. Thus, although Kant, like Moore, firmly separates intrinsic value from a thing's being desired for its own sake, he has resources for saying that a thing is objectively good as an end *because* it is desired for its own sake. And most things that are good will in fact be good in this way: they will be good because they are part of the happiness of a deserving human being.

On Kant's theory, the goodness of most things is, in the way described by Ross, relational—relative to the desires and interests of people. But since it must also be appropriately related to one thing that has intrinsic value, it is not merely "subjective." Value does, in Ross's extravagant terms, "shine with a reflected glory," and it is "borrowed rather than owned" by most of the things that have it. But it does have an original source that brings it into the world—the value-conferring power of the good will.

V

In this section I want to focus on some advantages of the Kantian way of describing values. In the next section I will show how some of these advantages are shared by Moore. In the last section I will discuss what I take to be the most important advantage that Kant's theory of goodness has over Moore's.

Kant's treatment of the two distinctions and the relations between them allows us to describe certain kinds of everyday matters of value in a way that is more flexible and that I think is more natural than is available to us if the two distinctions are conflated or equated. This is especially so for certain cases of what we might call "mixed" values. I have in mind a variety of different mixtures. Take some examples: a luxurious instrument; a malicious pleasure; an unenjoyed exercise of one's higher faculties; or an undisplayed art object. Now the idea that a thing can have value under a condition, when combined with the reminder that instrumentality or usefulness is not the only possible condition (that is,

some extrinsically good things are valued as ends), will help us to describe such cases.

Consider, for instance, a common symbol of aspiration—a mink coat. Is it valuable as a means or as an end? One hardly wants to say that it is valuable only as a means, to keep the cold out. The people who want mink coats are not willing to exchange them for plastic parkas, if those are better protection against the elements. A mink coat can be valued the way we value things for their own sakes: a person might put it on a list of the things he always wanted, or aspire to have some day, right alongside adventure, travel, or peace of mind. Yet it is also odd to say it is valued simply for its own sake. A coat is essentially instrumental: were it not for the ways in which human beings respond to cold, we would not care about them or ever think about them. To say that the coat is intrinsically or unconditionally valuable is absurd: its value is dependent upon an enormously complicated set of conditions, physiological, economic, and symbolic. Certainly, it does not pass Moore's isolation test, so far as I can see. A universe consisting of a mink coat or of someone's having one, without the associations that can only be provided by the particular relations and causal connections under which we live, is not really imaginable, much less valuable. What would a coat *be*? It seems hard even to apply the isolation test here, for one is tempted to say that its instrumentality *is* one of the elements in the "intrinsic nature" of a coat, even though it can hardly be said to be a property the coat would have under any set of laws of nature. If its instrumentality is not one of its intrinsic properties, then one is regarding the coat as something else—an animal skin sewed into a peculiar shape, perhaps. But then it seems as if one must strip away the practically relevant properties of the coat in order to ask about its intrinsic value—and that cannot be right. It is equally absurd to say of such a thing that it is a mere instrument, just because its value is conditioned. The Kantian distinctions allow us to say that the coat is valued in part for its own sake, although only under certain conditions. It even allows us to say of certain kinds of things, such as luxurious instruments, that they are valued for their own sakes under the condition of their usefulness. Mink coats and handsome china and gorgeously enameled frying pans are all things that human beings might choose partly for their own sakes under the condition of their instrumentality: that is, given the role such things play in our lives.

Another possible advantage is that the independent use of the two distinctions will provide us with a way of talking about the relation of pleasure, enjoyment, and appreciation to other kinds of value that does not turn these mental states into ends to which everything else is a means. Activities of various kinds might be thought to be good under the condition that we enjoy them and not good at all for those who, for one reason or another, cannot enjoy them, without forcing the conclusion that it is only for the sake of the enjoyment that they are valued. Certain difficulties concerning the "higher pleasures" described by Mill or those activities that Aristotle says are "pleasant in their own nature" although not necessarily "to a particular person" might be dealt with in this way. But this is a suggestion I cannot pursue here.

Consider also the example of an extraordinarily beautiful painting unsuspectedly locked up, perhaps permanently, in a closet. Now a beautiful painting, I am supposing, is valued for its own sake. If the two distinctions are equated, we must say it has intrinsic value. Yet it is locked in a closet, utterly unseen, and no one is the better for its existence. Consider Moore's isolation test: is a universe with such a painting locked up somewhere intuitively better than one without it? Is a universe consisting of such a painting better than a universe consisting of something quite plain, with no viewers in either? These are curious puzzles: and Moore's isolation test seems to force us to ask the metaphysical-sounding question whether the painting has this property, intrinsic value, or not. Yet we *know* what the practically relevant property of the painting is: it is its beauty. Now on the Kantian type of account we can say that the painting is valuable for its own sake, yet so long as it remains locked up and unseen, it is no good at all. The condition of its goodness—the condition of the goodness of its beauty—is not met. That condition is that the painting be viewed. Yet although its value is not intrinsic, the painting may be objectively good for its own sake. If it were viewed, and the viewer were enraptured, or satisfied, or instructed by its loveliness, then the painting would be an objectively good thing: for the world would be, really, a better place for it: it would be a substantive contribution to the actual sum of goodness of the world. Notice, too, that this does not in the least mean that we have to say that the painting is only valued as a means to the experiences of appreciation. Those experiences are not an

end to which the painting is a means, but the condition under which its value as an end is realized.

I am not suggesting that the Kantian treatment solves all the difficulties in our thinking about these things, but only that it does not drive us immediately to the conclusion that all of these things, valued only under conditions and only in a network of relations, must be mere instruments or contributors to some further thing—pleasure or some "mental" state, which supposedly has the real value. The conflation of the two distinctions does tend to have this effect. In particular, when conflation leads us to the conclusion that a thing can only be valued as an end when it is intrinsically valuable, or valuable independently of all conditions and relations, we find ourselves led inevitably to the curious conclusion to which modern moral philosophers are indeed frequently led—that everything good as an end must be something mental, some kind of experience. I have already mentioned one line of argument that leads to this conclusion: some sort of experience, such as pleasure, seems to be a condition of goodness of so many good things. Another line of thought that leads this way is this: no matter how much the philosopher wants to insist that the value of a good thing must be intrinsic and so nonrelational, the sense remains that the goodness of a good thing must have something to do with its goodness *for us.* It cannot merely be a property, metaphysical and simple, which we perceive in things and respond to in an extraordinary way. So the fact that goodness must lie in some relation to human beings, evidently at odds with the theory that goodness must be entirely nonrelational, is dealt with by making goodness a property of something belonging directly to the human being—our experiences or states of mind. By making goodness lie in the experiences themselves, the philosopher rids us of the worry: but what if no one is around to care about this good thing? What good is it then? Kant's way of looking at it, on the other hand, enables us to explain why ordinary good things are good only in virtue of the fact that people are around to care about them without tempting us to the conclusion that the only good things are mental states and experiences.

To some, it may seem paradoxical to claim that things are good because we desire or choose them, rather than to say that we desire or choose them because they are good. Ross, for example, finds it clear that when we call something good we are thinking of it as

having some attribute, not as an object of interest; he thinks of our interest as inspired by the perception of the thing's goodness. We choose the thing because it is good. This picture is part of what gives power to the theory that goodness is not relative to interest, and of course there is a way in which it is true. For instance, when we want a certain kind of thing, we usually want one with the virtues of that kind of thing. And it is also true that what makes a thing a good kind of thing is its virtues. In this sense our choice may be called forth by a thing's goodness, rather than the thing's being good because of our choice. But when we inquire into the basis for calling certain properties of a thing its "virtues," we always come back to something that is relative to certain conditions of human life. It is our interests and the bases of our interests that make certain qualities *virtues;* so these facts cannot make goodness a nonrelational attribute.

This shows up most clearly in the everyday kind of case of "mixed" value, in which the distinction between what we value for its own sake and what we value for the sake of something else is itself overstrained. Take this case: there are instrumental reasons, good ones, for eating. It keeps you alive. But most people could not really be said to eat in order to stay alive. Certainly, only someone who didn't enjoy eating, perhaps because of illness or some damage to the taste buds, would say that he ate "in order to stay alive." Are we then to say that eating is an activity that also has an *intrinsic* value? (Perhaps then we should be glad that we are so constituted that it is necessary for us?) Or shall we say that people eat for the sake of enjoyment—that pleasure is an end to which eating is a means? Of course, you cannot exchange another pleasure for it; hunger pains will prevent that. Perhaps then we should say we eat as a means, not to obtain life and health, but just to avoid pain. Now the philosopher wants to say: the real end is painlessness. But again, only someone in a particular situation would say that. Is this then a complicated case, to which the ends of life and health, enjoyment, and painlessness contribute in various ways? And if this is a complicated case, where are we going to find a simple one? It is easier to say that food is a good thing under the condition that you are hungry—or rather, under the set of physiological and psychological conditions that make it both necessary and pleasant for human beings to eat. Those conditions determine what the virtues

of a good meal are, and not all of these virtues are instrumental properties. But this does not mean that you choose the meal in response to a perception of its intrinsic value, or of the intrinsic value of eating it. The conditions of our lives make various things valuable to us in various ways, and it is sometimes artificial to worry about whether we value those things as means or as ends. It is the conditions themselves that make the things good, that provide the various reasons for their goodness. The question is not whether the thing possesses a special attribute, but whether these reasons are sufficient to establish the goodness of the thing.

This point can be sharpened if we distinguish between the initial condition that makes an object a candidate for choice and the full complement of conditions that, when met, renders the thing good. In the cases under discussion in this paper, the initial condition is the thing's desirability as an end (or at least not merely as a means). I have tried to show that the sense in which we can be said to desire things because they are good—i.e., for their virtues—does not show that a desirable thing need have a nonrelational property of goodness. What we call virtues just are the features of the thing that, given our constitution and situation, we find appealing or interesting or satisfying to our needs. It remains just as true, as far as this goes, to say that the thing is good because we desire it as to say that we desire it because it is good. For its virtues are still relative to our desires, or, more accurately, to the conditions that give rise to those desires. The reason that one cannot, on a Kantian account, rest with the perhaps less paradoxical formulation that value is conferred by desire is that desire is not by itself a *sufficient* condition of the goodness of its object. This is shown initially by the sort of case in which one has a desire which one would be better off without. Short of endorsing Kant's view that "the inclinations themselves, as sources of needs, however, are so lacking in absolute worth that the universal wish of every rational being must be to free himself completely from them" (G 46/428), we can agree that there are desires that conflict with one's health or happiness or that are self-destructive or pathological or simply burdensome out of all proportion to any gratification their fulfillment can provide. This already shows that the existence of a desire is not by itself a sufficient reason for the realization of its object; further conditions exist. The criterion that reasons be universalizable will also, on Kant's ac-

count, limit the capacity of desires to serve as reasons and so to confer value. But although desirability is not a sufficient condition of goodness, it is still the initial condition of the goodness of many good things, and so a main source of the goodness of those things.[19] On the Kantian view, not everything valued as an end need be intrinsically valuable or self-justifying for there to be a sufficient reason for it. A conditionally valuable thing can still be fully justified, if the unconditioned condition of its goodness is met. Things that are not self-justifying can be justified by something else. In particular, ends whose condition is their desirability can be justified by the rational choices of human beings.

VI

But I have not meant to suggest that Moore himself is prey to all of the difficulties that arise when the two distinctions are conflated. Moore has his own way of dealing with these issues of "mixed" value, a problem in which he was keenly interested. In order to handle cases of mixed value, Moore introduced a device which he regarded as one of his best discoveries: the theory of organic unities. The theory of organic unities involves two important points. First, it turns out that intrinsic value, on Moore's account, usually belongs to "organic" wholes or complexes of certain kinds, not to simple things. Second, it is true of such a complex whole that its value "bears no regular proportion to the sum of the values of its parts":

> It is certain that a good thing may exist in such a relation to another good thing that the value of the whole thus formed is immensely greater than the sum of the values of the two good things. It is certain that a whole formed of a good thing and an indifferent thing may have immensely greater value than that good thing itself possesses. It is certain that two bad things or a bad thing and an indifferent thing may form a whole much worse than the sum of badness of its parts. And it seems as if indifferent things may also be the sole constituents of a whole which has great value, either positive or negative.[20]

[19] I would like to thank the Editors for prompting me to clarify the roles of desire and choice in conferring value.

[20] *Principia Ethica*, pp. 27–28.

In his last chapter, "The Ideal," Moore provides various examples. For instance: the mere existence of what is beautiful has *some* intrinsic value, but so little as to be negligible, compared to the consciousness of beauty. If the consciousness of beauty is taken to be the cognition of beauty, then it in turn is made much more valuable if accompanied by an appropriate emotional response, which Moore identifies with the appreciation. Yet appreciation of beauty is not an end to which the beautiful object is a mere means. If this were so it would not matter whether the appreciation were produced in us by something genuinely beautiful or not, and it does: appreciating something that is ugly may be bad. Instead of saying that the value of the appreciation is conditional upon its appropriateness, as one might expect, Moore says that the great intrinsic value of appreciating beauty does not belong either to the object or the appreciative state but only to the complex whole formed of both. But the goodness of the whole is not the sum of the value of the parts. For we have seen that the value of the beautiful object by itself is quite small, and the value of the appreciation, in another context, can be absolutely negative. Moore has similar things to say about his other cases: for instance love itself is a good thing, but if your beloved is a good person, the whole is better by *more* than the addition of your beloved's goodness. These conclusions are arrived at by the method of isolation: we compare the value of various isolated wholes, with and without the relevant element. The important thing is to avoid the mistake of thinking that the element itself possesses all of the value of the difference its presence makes. It was because of this mistake that the Greeks attributed intrinsic value to knowledge. Moore explains that really, knowledge by itself has little or no value, but that it "is an absolutely essential constituent in the highest goods, and contributes immensely to their value."[21] Similarly, the great value that has been placed upon pleasure, and the delusion that pleasure is the sole good, is attributed to the fact that:

> Pleasure does seem to be a necessary constituent of most valuable wholes; and since the other constituents, into which we may analyse

[21] *Ibid.*, p. 199.

> them, may easily seem not to have any value, it is natural to suppose that all the value belongs to pleasure.[22]

Indeed, getting the right account of the relation of pleasure to other sorts of value seems to have been one of Moore's major motives in introducing the idea of organic unities. Things like pleasure and knowledge have what the tradition has called "contributive value."

I hope it is evident from these examples that the principle of organic unities is meant to do the same job that the notion of a conditioned good in Kant's theory does: it allows us to say of certain things that they are valuable only under certain circumstances, or valuable only when certain other things are true or present, without forcing us to say that these kinds of things must be valuable merely as instruments. Contributive value takes on the role that conditional value plays in Kant's view. The remaining difference is that Moore makes no distinction between what would be in Kant's terms really unconditionally (or intrinsically) good and what would be objectively good as an end.

VII

The principle of organic unities is crucial to Moore, for it enables him to make some of the same distinctions and judgments that the Kantian divisions make possible. Like the Kantian distinctions, it gives us a more flexible way of talking about the value of everyday things; and like the Kantian distinctions, it makes it possible for us to explain the conditional character of a good thing without rendering that good thing a mere means. Moore, who separates pleasure from the consciousness of pleasure, even complains in one passage that if pleasure were the sole intrinsic good, consciousness would have to be regarded as a means to it.[23] But the principle of organic unities is also in a certain way perverse. The seeming difficulties that it solves in fact arise from the relational or conditional character of the value of most of the things that human beings

[22] *Ibid.*, p. 93.
[23] *Ibid.*, p. 89.

regard as good. Yet it is precisely this relational character that Moore, with his insistence on intrinsic value, wants to deny.

Suppose someone said: on Kant's view happiness is a conditioned or extrinsic value and the good will its condition. But the happiness of a good person is, on Kant's view, always good, good under any and all circumstances, for its condition is met. So couldn't we say of this, as well as of the good will, that it is intrinsically valuable? What this would amount to would be constructing an organic unity out of happiness and the good will, and showing that on Moore's account it has intrinsic value. Then the Kantian notion of "objective value" and Moore's "intrinsic value" are not so different after all.[24]

And the answer to that is that there is still a difference. For Moore's view, and the intuitionistic method of isolation, veil or obscure the internal relations within the organic unity in virtue of which the organic unity has its value. Whereas the Kantian account, which focuses on rather than ignoring the internal relations of the valuable whole, allows us to see why happiness is valuable in just this case and not in another case. Moore can only say that the combination of happiness and good will works (is a good recipe, so to speak) while happiness plus the bad will does not. Kant can say that happiness in the one case is good because the condition under which it is fully justified has been met (roughly, because its having been decently pursued makes it deserved). Those internal relations reveal the *reasons* for our views about what is valuable, while Moore's view tends to cover up these reasons. And this might be true in other cases as well: if we think that aesthetic response is only valuable when the object responded to is genuinely beautiful, or that friendship is only valuable when your friend is good, or even if we think that aesthetic response and friendship are just *more* valuable in these cases, then this has something to do with the reasons we think these kinds of things are valuable at all. On Moore's account the only relation in which the elements of an organic whole stand to each other is the relation of being elements in a single organic whole. They are all on a footing with one another. But if Kant is right there is an order within "valuable wholes," a condi-

[24]Ross does something very like this in his discussion of the relation of virtue and pleasure in *The Right and The Good*, pp. 135 ff.

tioning of some elements by others, that is hidden by treating these elements as just so many *ingredients*. This order reflects the reason why the wholes are good.

Another way to put the point is this: Moore's theory drives a wedge between the reason why we care about something and the reason why it is good. Or rather, since on Moore's theory it is a mistake to talk about why something is good, we should say that it drives a wedge between our natural interest in something and our moral interest in it. On Moore's theory if you say that the reason something is good is because someone cares about it, that could only mean that the person's interest was an element in an organic whole which had intrinsic value. But according to Moore the question why such a whole has intrinsic value must not be raised: it just has the property of intrinsic value; there is no reason why it has that property.[25] Yet it is because it has intrinsic value that we ought to make it an end in our actions. A thing's goodness becomes a property that we intuit and respond to in a way that seems curiously divorced from our natural interests.

The interesting thing about that is that Moore took up the idea of intrinsic value because he saw that objectivity was not all that we wanted from a theory of value. He was certainly right to think that the same people who are discouraged by subjectivism are discouraged by an evolutionary theory or others of that kind. But to me it seems that this discouragement has to do with the way in which such theories undermine the nature of our concern for the good. For instance, if goodness is mere fitness to survive, then the only way goodness matters is the way the biological survival of the species matters—and that doesn't cover everything we feel about the importance of living a good life. But what is the nature of our concern for intrinsic values as Moore describes them? Moore seems to find it obvious that when we have determined what is intrinsically good we shall have an interest in bringing that into the world. His anti-naturalistic arguments prevent him from giving any account of why this nonnatural property should be so appealing to us. Of course, the isolation test by which intrinsic values are discerned guarantees that we will only attribute them to something that appeals to us. But that does not provide a justification of our

[25]See *Principia Ethica*, pp. 142–44.

interest in the intrinsically valuable or even a motivational explanation of it. On the Kantian account, by contrast, the good end is the object of a rational choice. The things that we want, need, care for, are good so long as certain conditions of rational choice are met. Thus, the reasons that things are good bear a definite relation to the reasons we have for caring about them.

The primary advantage of the Kantian theory of goodness is that it gives an account of the "objectivity" of goodness that does not involve assigning some sort of property to all good things. Good things are good in the way that Ross describes as relational, because of attitudes taken up towards them or because of other physical or psychological conditions that make them important to us. Only one thing—the good will itself—is assigned an intrinsic value or inner worth, and even the argument for that is not ontological. If we regard ourselves as having the power to justify our ends, the argument says, we must regard ourselves as having an inner worth—and we must treat others who can also place value on their ends in virtue of their humanity as having the same inner worth.

If human beings have an intrinsic value by virtue of the capacity for valuing things, then human beings bring goodness into the world. The distinction between a thing that is intrinsically good and a thing that is extrinsically good yet valuable as an end allows for the possibility that the things that are important to us have an objective value, yet have that value because they are important to us. Objective goodness is not a mysterious ontological attribute. The things that are important to us can be good: good because of our desires and interests and loves and because of the physiological, psychological, economic, historical, symbolic and other conditions under which human beings live.

University of California—Santa Barbara