

# 1

## The Relational Nature of the Good

*Christine M. Korsgaard*

A machine is inanimate and passive, but we are agents. Our constitution is put in our own power.<sup>1</sup>

Joseph Butler

### 1. MOORE ON GOOD AND GOOD-FOR

G. E. Moore, always ready to volunteer when a straw man would otherwise be wanted, asserted the priority of the idea of ‘good’ over that of ‘good-for’ (here I mean as in good-for *someone*) as an ‘obvious’ truth. In *Principia Ethica*, he wrote:

In what sense can a thing be good *for me*? It is obvious, if we reflect, that the only thing which can belong to me, which can be mine, is something which is good, and not the fact that it is good. When, therefore, I talk of anything I get as ‘my own good,’ I must mean either that the thing I get is good, or that my possessing it is good.<sup>2</sup>

In Moore’s eyes, this point provided decisive grounds for the refutation of rational egoism, for, as he continues:

In both cases it is only the thing or the possession of it which is *mine*, and not the *goodness* of that thing or that possession. There is no longer any meaning in attaching the ‘my’ to our predicate, and saying: The possession of this *by me* is *my* good.<sup>3</sup>

<sup>1</sup> Joseph Butler, *Five Sermons Preached at the Rolls Chapel and A Dissertation upon the Nature of Virtue*, p. 15.

<sup>2</sup> G. E. Moore, *Principia Ethica*, p. 98.

<sup>3</sup> *Principia*, p. 98. I am not sure why Moore says in this part of the quotation that the possession of something can be mine. In the first part of the quotation, he says, more correctly as it seems to me, that it is only the *thing* that can be mine. The fact that I possess something seems to belong equally to everyone, if it has an owner at all.

There are some ways we might try to rescue the notion of ‘one’s own good’ within the context of Moore’s theory, but they turn out to give rise to problems. For instance, you can be a selfish lout, and want good things *only* for *yourself*; or you can be a greedy lout, and want *all* the good things for yourself. But these attitudes, on Moore’s account, cannot possibly be rational. On Moore’s account, if you believe that it would be good for you to have something, then you must believe that everyone has the same reason to promote the state of affairs in which you have it as you do. That also means that if you believe it would be good for someone *else* to have something, then you must believe that *you* have the same reason to promote the state of affairs in which he has it as he does. So it is rational to care *only* about getting good things for yourself if you actually believe that you are the only one who should have them. And you can believe that the state of affairs in which you get *all* of the good things for yourself is a good state of affairs only if you can actually believe that everyone else has the same reason to promote it as you do, which is plainly insane. Nor can we eliminate the insanity by universalizing, since we cannot consistently believe that we have reason to promote the impossible state of affairs in which each person has all of the good things.<sup>4</sup>

And, on second thoughts, we might question whether you can even *be* a selfish or a greedy lout. An intrinsically good thing, according to Moore, is what he called an ‘organic unity’, because goodness supervenes on complex sets of facts. Moore was eventually driven to the view that the complex sets of facts on which it supervenes include the people (or other sentient beings) who are, as we should more ordinarily say, enjoying the good thing in question. Moore thought that we identify intrinsic goods by the test of isolation. This means that something is intrinsically good only if it is the case that it would be good even if it were the only thing in the universe, showing that its intrinsic properties alone are sufficient to render it good. But intuition baulks at ascribing intrinsic goodness to insensate objects which are unappreciated or unenjoyed by any conscious beings who are capable of appreciating or enjoying them. So, what is good turns out to be, for example, not just this beautiful sunset, but someone observing this beautiful sunset with a proper attitude of awed

<sup>4</sup> The friends of atomistic reasons may respond by suggesting that of course we all do have a reason to promote each state of affairs in which someone has all the good things, since each such state will be gratifying to someone, although my reason for wanting you to have all the good things counts in favour of a different course of action than my reason for wanting me to have all the good things. So I will stipulate that we are talking about states of affairs that we judge to be good all things considered, for which the reason is decisive.

appreciation.<sup>5</sup> But then it becomes a little unclear what it even *means* to be a selfish or a greedy lout. We might try defining the selfish lout as someone who wants only those good states of affairs in which *he* is a participant to be realized. But on a little reflection, we can see that this will not do, for that set includes—just for example—all of those good states of affairs in which the selfish lout is justly punished for his selfishness, or humbled for his pride, to the satisfaction of all right-thinking observers.<sup>6</sup> Nor can we limit the relevant set of states of affairs to those in which his role is a *good one for him* without scrapping this whole line of thought and starting over.<sup>7</sup> As for the greedy lout, who wants all of the good states of affairs for himself: we cannot even conceptualize *him* now. For as Moore said in the passage I quoted a moment ago, ‘there is no longer any meaning in attaching the “my” to our predicate’. Good states of affairs get realized or not, but there is no sense in which they belong to anyone.

I have been ragging Moore a little, but there is an important lesson suggested by these thoughts. You cannot start with the notion of ‘good’, and construct the notion of ‘good-for’ out of it. At least if values are essentially intrinsic in Moore’s sense, then ‘good-for’, in the sense of *good-for someone in particular*, is not merely a derivative notion. Rather, it does not mean anything at all.

That seems counterintuitive. But is it only a problem about Moore’s own account? Although it will take some preparation before I can make the argument, later I will suggest that it is not.<sup>8</sup> So suppose we try starting at the

<sup>5</sup> Then why does Moore talk about ‘getting’ good things at all? Generally, Moore thinks we can use the term ‘good’ for partial goods, and things that are frequently elements in intrinsically good wholes. But in fact Moore’s views about the possible *intrinsic* value of unappreciated things evolved. When he wrote *Principia*, his own intuitions did *not* baulk at the idea that unappreciated beauty, for example, has value (see pp. 83–84). By the time he wrote *Ethics*, however, Moore believed that conscious states formed part of all intrinsically good organic unities (p. 70). This would seem to strengthen his view that, when we are not talking loosely, good states of affairs never ‘belong’ to people; they just have people (or perhaps other sentient beings) among their elements. Yet, in *Ethics*, Moore takes up questions about whether ‘an action that most promotes the general good *also always* promotes the agent’s own good’ as if this were still a substantive and intelligible question (see pp. 98–100). I am not sure what he means, at this point, by an agent’s ‘own good’.

<sup>6</sup> Perhaps Moore could meet this objection by denying that a universe that consisted only of a selfish agent being justly punished for his selfishness, and another agent looking on and properly appreciating that fact, would be a good thing. Despite the case I make in the text, I am not prepared to put up a fight about that one.

<sup>7</sup> See Korsgaard, ‘The Myth of Egoism’, pp. 93–95, for a similar argument.

<sup>8</sup> Just as a preview: I argue that if ‘for’ means ‘experienced by’ and ‘experienced by’ means something like ‘observed by’, that does not capture the wanted sense of ‘good-for’ either, since I can observe good things the benefit of which accrues to others, or from

other end, and making the concept of ‘good-for’ prior to the concept of ‘good’? This will also seem counterintuitive to some of you. If, as someone might rashly suppose, what is good-for someone is his getting something good, then how are we to identify *which* things are good-for someone, without first identifying which things are good?<sup>9,10</sup> Indeed, the very expression, ‘good-for’, seems to suggest that the good comes first, and then stands in some sort of a relation, the *for-ness* relation, whatever that may be, to the person or other animal for whom it is good.

For all that, however, I think that the notion of ‘good-for’ is the prior notion. Or, to put it a better way, I think there is something essentially relational about the notion of the good itself. I think that good means something roughly in the neighbourhood of ‘welcome’. Of course you never know with philosophers, but in saying that, I am hoping that none of you will be inclined to insist that things can be intrinsically and non-relationally welcome. I think there is such a thing as the good, only because there are creatures *for* whom things can be good; that is, creatures who can welcome or reject the things that they experience. In fact, I think that the idea of something’s being good without its being good-for someone should be

which I am excluded. I can observe another’s happiness, say. It is of no use replying that experiencing someone else’s happiness is not a way of experiencing its goodness, since then we must say what the difference between experiencing happiness and experiencing its goodness is. If you are tempted to say that in order to experience the goodness of happiness you must take pleasure in it, you are either transferring the problem to pleasure—just being conscious of *it* does not make it mine in the right way either—or agreeing with me that its goodness is relational. All of this will be clearer later on. The general point is that if ‘good-for’ is derived from ‘good’ we must be able to say what the ‘for-ness’ relation consists in. If it is not that the agent participates in the good and not that he is conscious of the good, it is hard to see what it might be.

<sup>9</sup> In fact, we seem to identify what is good-for a particular person, in a rough and ready way, by identifying which things are good-for anyone in a certain position, or of a certain kind, and then applying that conclusion—always defeasibly—to the person or to the animal at hand. This is suggested by Rawls’s account of the good, which I will be discussing later. I think the fact that we do this is important, for a reason that I will not be discussing in this chapter. The reason is that there are certain conceptual pressures in favour of identifying what is good-for X in terms of the species of which X is a member. The issue is this: the good-for X is relative to X’s nature, but we do sometimes say that X would be better off if X changed his nature. For instance, a person might be better off if he had more refined aesthetic tastes. But then a question arises about how much and what sort of change to X’s nature we can be talking about, and still be talking about X. Is there any sense in asking whether the squirrel I now see outside of my study window would be better off if she were Aristotle, or had Aristotle’s appetite for knowledge, for example? For various reasons, X’s species seems like a plausible place to draw the line.

<sup>10</sup> When I say that it is rash to suppose that what is good-for X is his getting something good, I have in mind cases like this: Aristotle says that the contemplative life is best, but *for* some people, the political life is better. That is a complicated thought however you twist and turn it, but I do not see *any* way to get there if you simply suppose that what is good-for someone is simply getting something good.

rejected as unintelligible. So in this chapter I hope to defend the relational nature of the good. I will also show how an understanding of the essentially relational nature of the good paves the way for what I call a ‘constructivist’ account of the good.<sup>11</sup>

## 2. TWO KINDS OF GOODNESS

In order to clarify the thesis I am defending, I need to make some distinctions. This will also give me an opportunity to explain the relationship between my thesis and some other things that philosophers have been saying lately about the good. We tend to use the concept of the good in two different ways, and the relation between them, as I have argued elsewhere, is a little unclear.<sup>12</sup> First of all, ‘good’ is our most general term of evaluation. Nearly any kind of thing, certainly anything we have a use for, or interact with, can be characterized as good or bad. We characterize machines and instruments, dogs, cats, and people, food and weather, jobs and schools, and myriad other things as good or bad. I will call that good in the ‘evaluative’ sense. Second, we speak of something we call ‘the good’ or (interestingly) ‘the human good’, or (traditionally) the ‘*summum bonum*’, which is supposed to be the end and aim of all our strivings, or the crown of their success, the ultimate thing that we want or ought to want to achieve or obtain. I will call that good in the ‘final’ sense. Philosophers have had oddly little to say about the difference between these two uses of the concept of good, and what the relationship between them might be. Why do we even use the same term for both? When we use the term ‘good’ evaluatively, we

<sup>11</sup> A more complete defence would require addressing two objections that I do not have space to discuss here.

First, the view apparently makes aggregation impossible, since what is good-for me plus what is good-for you is not good-for anyone in particular. In so far as we have some intuitions in favour of aggregation, we need to ask whether they are valid and if so what they are based on. I have briefly discussed that problem in ‘Interacting with Animals’, p. 96. The account I give there suggests that some of the actions whose rightness we explain by appeal to aggregation might better be explained by appeal to the idea of making Pareto optimal moves. The view I defend in the last section of this chapter lends support to that thought, for moves that are Pareto optimal are, as Rawls pointed out in *A Theory of Justice*, better-for us all—they are shareable goods.

Second, the view disallows certain claims of the form ‘it is good that...’ or ‘it is better that...’ that seem intuitively possible to make, such as the claim that it is better that the universe should contain some intelligent life forms than that it should contain none at all. Assuming that the universe is not the kind of thing for whom things can be good or bad, we need to ask what, if anything, we are to make of such claims. I say a little about how to address these difficulties in ‘On Having a Good’.

<sup>12</sup> I make the case in ‘The Origin of the Good and Our Animal Nature’, which covers many of the same ideas discussed in this chapter, but from a slightly different angle.

seem to be referring to the thing's performance of or its capacity to perform in its role or function, in a very broad sense of role or function. A good thing in the evaluative sense, as philosophers since Plato and Aristotle have been telling us, is one that has the properties that enable it to perform its role or function successfully and well.<sup>13</sup> When we use the term 'good' in the final sense, however, it is less clear what property we are referring to, or what standard of evaluation we are invoking. I think that the reason why so few philosophers have remarked on this point is that they have assumed that when we talk about the final good, we are still using the term evaluatively. We are evaluating a person's whole life, or the ends she pursues, perhaps. Certainly we often say of someone who has died that she had a good life, with a slight suggestion, at least for the secular among us, that one can hardly ask for more. But how can we evaluate a life? A life does not, or anyway does not obviously, have a role or a function, so it cannot have the properties that enable it to perform its role or function well. So what do we mean by saying that a life was good? Sometimes we seem to mean simply that it had a lot of good things in it: the person had a good job and did it well, had a good marriage, raised happy, healthy children who loved her and did not disgrace her, enjoyed some aspects of nature and the arts, and so forth. But then it seems as if at least some of *those* things must be good in the final sense. How do we determine that? Such things are supposed to be good as ends, but how do we evaluate something's suitability to be an end? Do ends, considered as such, have a role or a function? To say that they are *ends* seems to imply precisely that we do not want them for anything else. But then how do we propose to evaluate them when we ask whether they are good? Or if we are not evaluating them, then what exactly are we doing when we say that they are good?<sup>14</sup>

I think we are saying something relational: that is, something that is essentially relative to the condition of some person or animal. So here is the promised clarification of my thesis. 'Good' in the evaluative sense is obviously relational. That part of my claim is not controversial. An evaluatively good thing is good-at or good-for some purpose or function. It seems natural to say that *what* evaluatively good things are good relative *to* is given by the final sense of good—by whether it is a good thing, in the context at hand, that the purpose or function of the evaluatively good thing should

<sup>13</sup> See Aristotle, *Nicomachean Ethics*, 1.7 1097b21ff., and Plato, *Republic*, Book I, 352d–354.

<sup>14</sup> An expressivist philosopher might be tempted to say we are commending them, or maybe recommending them. Then what the things that are good in the final sense and the things that are good in the evaluative sense have in common is that we recommend them. But that just postpones the problem. On what grounds do we recommend things as finally good? I will say more about commendation below.

be served. I am claiming that goodness in the final sense is also relational. So it can only be a good thing, in the final sense of good, if it is good-for someone in particular. Notice that a question now arises that parallels one I asked before. I asked how on Moore's theory we could get from something's being good to its being good-for someone, and I agreed with Moore that, at least on his conception of intrinsic value, we cannot. Can we get from something's being good-for-someone-in-particular to its being good simply? That is a question I will take up later, in section 7. For now, I want to argue that good in the final sense is in the first instance relational. It is not true that we need to know what is good before we know what is good-for someone, since despite its surface grammar, the notion of *good-for someone* is in fact the prior and more fundamental notion.

### 3. GEACH AND FOOT ON ATTRIBUTIVE GOODNESS

It will be helpful to preface my own argument by distinguishing it from one that makes a similar point. In his famous paper 'Good and Evil', Peter Geach argued that 'good' is an 'attributive' rather than a 'predicative' adjective.<sup>15</sup> An attributive adjective cannot be predicated of the noun it modifies, as it were, independently. We can say that a red convertible is red, so 'red' is predicative, but we cannot say that a large mouse is large, so 'large' is attributive. Good is attributive, Geach thinks, so it is true both that when you say that something is good you are always saying that it is a good X or Y, you are never just saying that it is good *simpliciter*, and also that when you say that something is a good X or Y, you are not necessarily suggesting that we should approve or welcome it, as the case of, say, 'good assassin' shows.

Geach's argument looks at first as if it is intended to deny the use of the concept of 'good' in what I have called 'the final sense' altogether. Although in this paper Geach does not mention the word's use as a noun, as when a philosopher talks about The Good or the *summum bonum*, he plainly thinks that calling something a 'good thing' or 'good event' would be about as close as we could come to calling it, quite simply, 'good'. (Philippa Foot later added 'good state of affairs' to this list.) Geach dismisses such expressions as illegitimate attempts to disguise predicative uses of the term.<sup>16</sup> He says:

'Event', like 'thing', is too empty a word to convey either a criterion of identity or a standard of goodness; to ask 'Is this a good or bad thing (to happen)?' is...useless...unless the emptiness of 'thing' or 'event' is filled up by a special context of

<sup>15</sup> Geach, 'Good and Evil', pp. 33–42.

<sup>16</sup> 'Good and Evil', p. 34.

utterance. Caesar's murder was a bad thing to happen to a living organism, a good fate for a man who wanted divine worship for himself, and again a good or bad act on the part of his murderers; to ask whether it was a good or bad event would be senseless.<sup>17</sup>

Geach's worry seems related to the one I voiced when I pointed out that 'good life' cannot obviously be an evaluation of a life, because evaluation refers to something's ability to fulfil its role or function, and life does not have a role or function. In the same way, we might understand Geach to be claiming that things and events, considered just as such, cannot be evaluated, because there is nothing things and events, considered just as such, are supposed to do, no function that they serve.<sup>18,19</sup> That is why it is tempting to say that Geach is rejecting the use of 'good' in the final sense altogether.

Admittedly, that would be a rather surprising move for such an admirer of Aquinas, who made much of the idea of a *summum bonum*. But in any case it is not what is happening. The idea of 'good-for' in the sense of good-for someone makes an appearance in Geach's remarks here: death is bad-for an organism, but being murdered in the way Caesar was is good-for a man who seeks to be worshipped as a god. So apparently Geach thinks that being good-for someone supports the attributive use in something like the same way that having a role or a function does. That means that Geach thinks there are two different ways in which 'good' can be attributive: speaking roughly, we say X is good-*for* P or sometimes good-*at* P when P is a purpose that X serves or a role that X carries out, and we also say X is good-*for* P when P is a person or an animal who is benefited by X.<sup>20, 21</sup> In other words, there is both a final and an evaluative sense of good-for. So the *summum bonum* will survive, but presumably as something that is good-for human beings considered just as such.

<sup>17</sup> 'Good and Evil', p. 41.

<sup>18</sup> Except possibly to be good, one feels tempted to say, but that only takes us around in circles.

<sup>19</sup> I say we 'might' understand Geach this way because Geach does not mention a functional criterion, for reasons that will become clear shortly.

<sup>20</sup> Philippa Foot also notes the connection between 'good-for' in the sense of good-for a person and 'benefited', in *Natural Goodness*, pp. 93ff.

<sup>21</sup> Actually, we do not often say that X is good-for P when P is a purpose that X was made to serve or regularly serves. There is little occasion for saying that knives are good-for cutting, since anyone who knows what a knife is already knows that. We are more likely to say it about unexpected uses—that rock would make a good rolling pin—than about regular ones. It is not exactly the same with the use of good-for as in good-for someone. We are more likely to eschew that use of good-for when the thing is not merely good-for the someone in question, but necessary. We do not say that air and water are good-for animals. The understatement makes such a remark sound like a joke. For that matter, *pace* Geach, we do not say death is bad for animals either, because of the understatement; although in a deviant case—perhaps when an animal is suffering uselessly—we might say death is good-for the animal.

Now that is a conclusion I am perfectly happy with. It suggests Geach would agree with me that good in the final sense is a relational notion—a form of good-for. But that is exactly the conclusion that I think needs a further defence here. For I also agree with Geach that when we say something is good, we should ‘convey a standard of goodness’. And I think that this second use of good-for—namely good-for some person or animal—does not carry with it the same straightforward standard of goodness that being *good-for* a purpose or *good-at* a role does. When we seek a standard for applying this notion, we will run into the question I mentioned earlier, namely, how we can say that some thing is good-for someone without appealing to some prior notion of what is good.<sup>22</sup>

One might suppose that by deploying Philippa Foot’s notion of ‘natural goodness’ we can make this problem go away, although Foot herself does not think this, as I will notice below. Foot observes that we can identify natural goodness and defect in the properties of organisms simply by considering how the organism carries on its activities and what it needs in order to do so. So for instance, there is no mystery about how we can say that stealth and swiftness are good properties in a tiger. And we might think it is just as obvious that these properties are good-for the tiger. A similar point seems in order when we think about someone as occupying a role. A steady aim and a stony heart are good properties in an assassin, and they are also good properties for you to have, in the sense of good-for you, if you are an assassin. Or anyway, that is true, *at least in so far as you identify with the role*. That is a qualification that I will come back to.

<sup>22</sup> Foot, in *Natural Goodness*, notes that we do not need to invoke claims of the form ‘it is good that . . .’ in order to talk about natural goodness and defect in plants and animals. In order to say that swiftness and stealth are good properties in a tiger, for example, we do not need to say that it is good—that the tiger should survive. I should note that the issue I am discussing here is orthogonal to that one. The question whether we can make claims about good-for in the sense of good-for a person without first making claims about which things are good is not the same as the question whether we need to make claims of the form ‘it is good that . . .’ before we can make such claims. To say that something is good is not the same as to say that it is good that it should exist. Moore’s view might mislead us about this, I think. On his view something is good if it would be good if it were the only thing that existed, but of course if it is *not* the only thing that exists, then it might not turn out to be good that it does exist, given the other things that do. Perhaps the beautiful sunset distracts the viewer from seeing something far more important, the tragedy that is unfolding at his feet. But that means if we put these additional circumstances into an organic unity with the original thing we get something bad, or anyway worse than the original thing by itself. This line of thought seems to push us towards the idea of one big organic unity, which would be good *simpliciter*, and of which we could also say that it is good that it exists. That makes the two notions, that of something’s being good, and that of its being good that something exists, look the same. But we only need to accept this view of ‘good-that’ if we accept Moore’s account of value, and especially the isolation test, in the first place.

This strategy for thinking about the human good is of course associated with Aristotle's proposal that we can discover the human good by determining the human function. Aristotle is confident that the properties that make us good-at being human, which he identifies with the virtues, will also be the properties that somehow constitute or guarantee our good.<sup>23</sup> Some latter-day virtue ethicists are also tempted by that thought. But that conclusion cannot be earned simply by *assimilating* good-for in the final sense to good-for (or good-at) in the evaluative sense without further ado. As Foot herself insists, 'It is too quick to say that because human *goodness* belongs to those who have the virtues, human *good* is what they will attain in acting well.'<sup>24,25</sup> How do we make that connection, if it can be done at all?

I already gestured at part of what I think is wanting when I said that having the properties that make you good-at being an assassin will also be good-for you *in so far as you identify with the role of assassin*. But, surely, you will protest, we can say that the properties that make a tiger good-at being a tiger are also good-for the tiger without assuming that the tiger identifies with her role? As I will argue later, there is a reason we are confident that what is good-for a tiger is to be good-at being a tiger, and it is not just that we ignore the difference between those two ideas. It is, more or less, that we know that the tiger cannot help but identify with her role.

That leads me to one last point about Geach. Geach thinks that his account of the grammar of good renders claims about the good straightforwardly descriptive, and on this basis he denies non-cognitivist accounts that claim that 'good' has a primarily commendatory force. These days it is a commonplace in the philosophical literature that when we use the term 'good' in the ordinary evaluative or functional sense, we can be straightforwardly descriptive. That a good assassin has a steady aim and a stony heart are things that we apparently can say without commendation. I do not think that is quite right. But the fact that we can apparently make the

<sup>23</sup> For my own view about how that works, see note 25.

<sup>24</sup> *Natural Goodness*, p. 92. Foot suggests that Elizabeth Anscombe and Gavin Lawrence are guilty of this sort of slide.

<sup>25</sup> Although I agree with Foot about this, I do not agree with her way of reading *The Nicomachean Ethics*. We cannot just describe human activities and then say, 'These are the things you need to be able to do in order to be a non-defective human being.' Elsewhere I have argued that what Aristotle meant by 'function' (*έργον*; *ergon*) is neither a purpose that something serves nor the activities in which a thing characteristically engages, but *how* a thing does whatever it does. The distinctively human way of doing things is through rational choice. Since a 'virtue' (*ἀρετή*; *arete*), in the strict sense in which Plato and Aristotle used the term, is a quality that makes you good-at your function, the moral virtues must make us good-at rational choice. *The Nicomachean Ethics*, I believe, should be read as an argument that the qualities that we ordinarily call 'the moral virtues' are virtues in this technical sense. I make the case in 'Aristotle on Function and Virtue' and 'Aristotle's Function Argument'.

connection between having the properties that make you good-at performing a role and having the properties that are good-for you to have by supposing that you identify with the role suggests a slightly different account of what the supposedly ‘non-cognitive’ element is. Judgements of good-for, in the sense of *good-for-someone*, are ones that incorporate a certain point of view. As we will see, this turns out to be—but for non-trivial reasons—the point of view of the someone for whom the things are good. The use of judgements of goodness in the practice of commendation is just an artefact of that. To put it another way, the judgement that something is *good-for* someone is essentially *empathetic*, or sympathetic in the Humean sense. It is a judgement we make essentially by taking up the person’s or animal’s own point of view, and so sharing in his evaluative responses.<sup>26</sup>

#### 4. RAWLS ON GOODNESS AS RATIONALITY

I now turn to a slightly different attempt to make a connection between the evaluative and the final senses of good. In *A Theory of Justice*, John Rawls argues that our criterion for applying the evaluative notion of good can be extended in a natural way to cover the case of final goods.<sup>27</sup> According to Rawls, to say that something is good is essentially to say that it has the properties that it is rational to want in that kind of thing. The everyday use of the term ‘good’ supposes that the person doing the wanting wants the thing for the reasons for which one usually wants such things. When we say that a certain make of car is a good one, for example, we mean that it has the properties it is rational to want in a car, given what cars are usually used for, say, a safe means of transport for people and their various forms of baggage. When we say that something is a good X for someone in particular, we can relativize the notion to any special interests that that particular person may have. So I can say that given that my brother travels on business and essentially lives out of his car when he does, a minivan is a good car for him, though I would not recommend one for everyone.

Rawls’s crucial move is to extend this kind of evaluation by defining the notion of a rational plan of life. I will not go into the details about how he

<sup>26</sup> As I will be mentioning later on, in ‘Realism and Constructivism in Twentieth-Century Moral Philosophy’, I argue that metaethical differences should be understood in terms of the function of moral concepts. So you might suppose I have just offered an expressivist account of the good, based on the claim that we use the concept of good primarily to express a certain state of mind, sympathy. Certainly we can use it that way, but I do not think that is the primary function of the concept, as I will explain later.

<sup>27</sup> Rawls, *A Theory of Justice*, ch. 7.

defines it, since they do not matter for the use I am making of Rawls's argument here. Suffice it to say that Rawls identifies what he takes to be some plausible rational principles for constructing, and choosing among, possible plans of life. The things that are rational for us to want are the things that are parts of, or contribute to, our rational plans of life.

The idea of something's being rational to want is helpful in this context because when we are using the term 'good' in the evaluative way, it captures the same content that the idea of something's role or function does. Since knives are ordinarily wanted for cutting, a good knife is a sharp one, and if you want a knife for the usual reasons, sharpness is a property it is rational to want in your knife. So instrumental or functional properties, broadly speaking, coincide with the properties it is rational to want. But, with Rawls's idea of the life plan in hand, we can extend the idea of 'rational to want' to a person's ends. Earlier, when I asked what the notion of 'good end' might mean, I asked whether ends have a role or a function, and I denied that they did. But, in effect, Rawls's move *does* assign our ends a kind of role or function. Their role or function is to serve as an element in a person's rational plan; some of them are better than others at playing that role.

But there's another way to characterize Rawls's move here that I think is even more important. I have now described the evaluative notion of the good in two different ways: first, as invoking the plain, descriptive idea that something has the properties that enable it to serve its function well; second, as invoking the slightly more normative idea that something has the properties it is rational to want in that kind of thing. The difference between these two ways of thinking of evaluative goodness is that when we think of the object in the slightly more normative sense of being rational to want, we consider its functional properties *from the point of view* of someone who wants that sort of thing. This enables Rawls to establish a continuity between the evaluative and the final good, since the final good as Rawls conceives it is also characterized from the point of view of the one whose good it is. That is, it is characterized as what it is rational *for* that person to want, given his rational plan. So both evaluative and final goodness are relational; they are goodness relative to someone's plan, and therefore goodness *for* that person.

Even the most general evaluative use of 'good', when it is not relativized to the point of view of any particular person, evokes a relation to a point of view. If I just say, for instance, 'the Honda is a good car', without any qualification, I mean, 'rational *for* pretty much anyone in want of a car to want, given what such things are generally used for'. That ties the goodness of the car to 'pretty-much-anyone-in-want-of-a-car's' point of view. In my view this is no accident, for the concept of the good always makes an essential reference to someone's point of view. In fact, that is putting it too mildly,

for as we will see there is only such a thing as the final good because there are beings who *have* points of view. That is why the final good is a relational concept, as I will now try to show.

## 5. HOW HEDONISTIC UTILITARIANISM WAS ALMOST TRUE

I am going to begin by offering an explanation for why there are such things as final goods and bads at all, why—if I may put it this way—the natural world we live in is one that contains states of affairs to which these normative concepts apply. My explanation, although not uncontroversial, is in one sense unsurprising. I believe that there are good and bad states of affairs because there exist in the world beings *for whom* things can be good or bad in a specific way. The beings in question are the ones who are sentient, or conscious—roughly speaking, the animals. The thesis is unsurprising because it is a thesis that is also held by some philosophers who defend a very different philosophical outlook from my own—namely, the hedonistic utilitarians.<sup>28</sup> In fact I think that the reason why hedonism is so perennially tempting is that the idea that the good is pleasure captures, or anyway wants to capture, the relational nature of the good. But hedonistic utilitarians promptly lose this advantage by making pleasure intrinsically rather than relationally good after all, in order to make the aggregation of goods across the boundaries between persons (or animals) possible. Relational goods cannot be added across such boundaries, since what is good-for me plus what is good-for you is not good-for anyone in particular.<sup>29,30</sup> I hope to make my position clear by explaining how it differs from that of the hedonistic utilitarian.

<sup>28</sup> Why don't I just say the hedonists? Because the fact that utilitarians think that pleasures can be added across the boundaries between animals shows something important about their conception of pleasure itself—that pleasure itself is not really relational—and I do not know that all hedonists conceive of pleasure in that way.

<sup>29</sup> As I explain in my discussion of aggregation in 'Interacting with Animals', I think it is possible that what is good-for me plus what is good-for you might be better for us, as long as we are only adding goods and not subtracting them from either of us. See note 11.

<sup>30</sup> Is there also a problem about adding relational goods across the boundaries between time-slices of persons, or of animals? Utilitarians like Parfit think that the idea that I can be compensated for a pain I experience now by a greater pleasure I experience later is just as problematic as the idea that I can be compensated for a pain I experience by a greater pleasure that you experience. See Parfit, *Reasons and Persons*, pp. 329–339, and my response in 'Personal Identity and the Unity of Agency: A Kantian Response to Parfit', especially pp. 384–386. See also note 53.

Hedonistic utilitarians believe that there is a certain conscious state, called pleasure, to which final or intrinsic value attaches, and another conscious state, pain, to which final or intrinsic disvalue attaches.<sup>31</sup> Conscious beings bring goodness and badness into the world because their consciousness gives them the capacity to realize these intrinsically good and bad states. The consciousness or sentience of a conscious or sentient being is, as it were, the place where the goodness and badness of the world occurs.

Four features of this view seem problematic, at least to me. I will discuss three of them in this section, and another later on. The first is familiar. The view seems to suggest that nothing can be good or bad for anyone except the character of his or her conscious experiences. On intuitive grounds, many philosophers have rejected this conclusion. What you do not know can hurt you, and not just because it is likely to intrude on your consciousness sooner or later. If everyone you suppose loves you actually hates you, that is a bad thing for you, and not just because you might find it out. Objective states of the world, and not just states of the mind, are good and bad for people and animals. So most of us think.

The second problem is that the account leaves in place a problem I mentioned at the beginning—the problem of the relation between the evaluative and final senses of good. It throws no light on what, if anything, the relation between these two ideas might be. It certainly throws no light on why we might be tempted to think that the two ideas are connected in the way that Aristotle thought: that we secure the human good by being good in the evaluative sense, or virtuous. Of course, we might think that Aristotle was just confused about that, but we should at least give an account of why one might be tempted to connect the two ideas in that way. If we accept the hedonistic utilitarian view, as I said earlier, it is not even clear why we use the same word.

Before I expound this point I should explain how I think the evaluative notion of goodness, the parallel to ‘virtue’ here, applies to non-human animals. There are two ways in which it does so, but only one of them is important here. We often describe an animal as being good or bad when the animal has properties that serve our own purposes. If you say of a certain kind of chicken that it makes a great roaster, or of a certain kind of horse that it is a good little saddle horse, or of a breed of dog that it makes a fine watchdog, you are talking about the animal as if it were, like a machine, designed to suit human purposes. In all these cases the animal is evaluatively good

<sup>31</sup> Obviously, utilitarians come in many different stripes and sizes, and not all of them will recognize themselves in the characterization I am about to give. It will be enough for my purposes if the reader recognizes what I say as describing a fairly central case of utilitarianism and is able on the basis of it to understand the contrast I wish to set up.

in the sense that it has properties that enable it to serve what *we* think of as its function well. In Rawlsian terms, the chicken's plumpness, the horse's cooperative nature, the dog's ferocity, are properties that it is rational for us to want in these animals, given what we use these animals for. This form of evaluative goodness is not a promising place to look for anything like an Aristotelian connection between evaluative and final goodness, since no one is tempted by the view that an animal's final good is somehow related to its evaluative goodness in this sense. No one thinks, for instance, that the good-for a chicken lies in being a good roaster.

But when we regard an animal as something that is in effect 'designed' by nature to survive and reproduce, then again we think of it as having functional properties, properties that enable it to serve 'nature's purposes' well. An animal that has such properties is essentially a healthy one, well suited to living a full life of its specific kind. These are the kinds of properties that Foot has in mind when she talks about natural goodness, like the stealth and swiftness of the tiger. This is a form of evaluative goodness, but it is not relative to our purposes. Should we say that it is related to the animal's own purposes? Interestingly, as I mentioned before, we *are* inclined to say that having these properties is good-for the tiger herself, as well as that they make for a good tiger, and so we do connect evaluative and final goodness in this case. But there is a question about why that is so.

The hedonistic utilitarian answers that question in a particular way. He says that the good-for a tiger, just as for anything whatever, is to have pleasant experiences. He posits a causal relation between the tiger's satisfaction of her needs, and quality of her experiences. If she hunts down enough food for herself and her cubs, she will feel satisfied; if not, she will be pained by her own hunger and distracted by the whining of her cubs. So there is a causal connection between the tiger's evaluative goodness in the sense of how good she is at being a tiger, and the goodness of her experiences. But it is only a causal connection, not a conceptual one; hooking the tiger up to a Nozickian experience machine, which would feed pleasant sensations directly into her brain, would be just as good-for her.<sup>32</sup> A causal connection between evaluative and final goodness of this type throws no light on the possibility of a conceptual connection between the two notions, or on the standard of evaluation we are invoking when we say that something is good in the final sense. The answer to the question of what we mean by saying that pleasure is the good—or rather, what standard of evaluation we are applying when we say that it is good—is just as elusive as it was before.

The third problem may feel like it is just a puzzle, but in fact I think it is the one that points the way to what has gone wrong here. The third problem

<sup>32</sup> Robert Nozick, *Anarchy, State, and Utopia*, pp. 42–45.

is the one suggested by Moore's argument. As we have already seen, Moore thinks that if you start from the notion of intrinsic value, at least in his sense, you lose any possibility of making sense of the notion of good-for. So the fact that the tiger is having an intrinsically valuable sensation does nothing to show that the sensation is in any special sense *good-for her*. She is merely an element in a state of affairs that is good: a tiger's having a pleasant sensation. We may be tempted to say that if the pleasures of a tiger are a good thing, that is only *because* they are a good thing *for* the tiger. But on Moore's conception, we cannot say that, since on Moore's conception of intrinsic value, goodness simply supervenes on certain complex states of affairs, and no reason for their goodness can be given.<sup>33</sup> So the tiger's having a pleasant sensation is just a good state of affairs in which she happens to be a participant, and is in no way good-for her.

Is the hedonistic utilitarian really stuck with this conclusion? The hedonistic utilitarian thinks that pleasant experience itself meets the isolation test, and so is a good thing. So perhaps we can say that when the tiger has a pleasant sensation she *has* something good, at least if we assume that experiencing something is a way of 'having' it. But can we? If experiencing something is just a matter of being conscious of it, that is, of knowing of it directly, then the mere fact that I experience something cannot possibly make it good-for me. After all, when I see my rival hand in hand with my beloved, I experience their intimacy and happiness, but I will hardly regard this as a boon. On the hedonistic utilitarian theory, this is the relation in which I stand to pleasure: it is a good thing of which I happen to be conscious.<sup>34</sup> Beyond that, it has nothing special to do with me.

What has gone wrong here is that the hedonistic utilitarian conceives of pleasure as an *object* of experience, rather than as a way we experience things; as something we are related to, rather than as a way we relate. I believe that the grammar of experience talk is part of what misleads us here. For example, colloquially, we can speak either of 'the experience of loss' or 'the experience of grief'. The parallel structure suggests that grief and loss are on a footing, two possible objects of experience. But of course they are not. Grief, or as one might better say, grievingly, is the *way* we experience loss. In my own view, grief is a form of *rejection*, and, as such, is the kind of pain we undergo when everything in us is in a state of massive rebellion against a personal loss.

To explain this remark I need to rehearse an argument that I made in *The Sources of Normativity*.<sup>35</sup> There are in fact two (or at least two) quite different

<sup>33</sup> See Korsgaard, 'Two Distinctions in Goodness', pp. 270–273.

<sup>34</sup> See note 8.

<sup>35</sup> Korsgaard, *The Sources of Normativity*, §§4.3.1–4.3.11, pp. 145–155.

conceptions of what pleasure and pain are. According to what I will call the Benthamite view, pleasure and pain are particular sensations, varying, as Bentham would have it, only in intensity and duration.<sup>36</sup> According to the other view, which I will call the Aristotelian view, pleasures and pains are not sensations, but reflexive reactions to the things we experience. Specifically, they are reactions to the objects of experience as welcome or unwelcome, as to-be-accepted and if possible to-be-continued in the case of pleasure and as to-be-fought-off and if possible to-be-stopped in the case of pain.

As the name I am giving it suggests, the second view has a philosophical heritage in the views about pleasure and pain put forth by Aristotle in *The Nicomachean Ethics*, especially in Book X. In the Book VII discussion, Aristotle identifies pleasure with activity,<sup>37</sup> while in the Book X discussion, he puzzles his readers by saying instead that it is something that ‘supervenes’ on activity.<sup>38</sup> He associates pain with something like obstacle or difficulty. Pain is the state you are in when the activity you are engaged in is too difficult, or too easy and therefore boring, or when you struggle to keep doing it although something outside is distracting you. Pleasure is the state you are in when you are wholly absorbed in your activity and want it to go on for ever.<sup>39</sup> It is an advantage of the Aristotelian view that it can explain the painfulness or pleasantness of things that are not necessarily accompanied by any particular sensation—not only activities but other forms of experience as well. Scraping your knee and breaking your heart are both painful; taking a hot shower and reading a great novel are both among the pleasures of life.

In understanding the difference between these two accounts, it is important not to read the explanatory ambitions of the Benthamite view into the

<sup>36</sup> The idea that there is some one thing that you feel whenever you are doing or undergoing something that you like, and some one thing that you feel whenever you are doing or undergoing something you do not like, seems out of keeping with experience, so some philosophers would prefer to make pleasure and pain each a family of sensations, with every pleasure similar to every other in respect of its pleasantness, and every pain similar to every other in respect of its painfulness. See, for just one example, Hume: ‘under the term pleasure, we comprehend sensations, which are very different from each other, and which have only such a distant resemblance, as is requisite to make them be express’d by the same abstract term’ (*A Treatise of Human Nature*, p. 472). I do not know how to make sense of this, though. Is the respect in which they are similar *itself* supposed to be a sensation, or a family of sensations? If having a sensation is something like performing a basic or uninterpreted act of discrimination, the idea that we then notice similarities and differences in our sensations seems to require further discriminations among our most basic acts of discrimination. But this gets me into large questions that I am not qualified to take on, so I will leave it alone.

<sup>37</sup> Aristotle, *Nicomachean Ethics*, 7.12 1153b14–15.

<sup>38</sup> *Nicomachean Ethics*, 10.4 1174b31–35.

<sup>39</sup> *Nicomachean Ethics*, 10.5.

Aristotelian view. On the Benthamite view, pain and pleasure are sensations that we experience when we engage in certain activities, and we appeal to the character of those sensations in order to *explain* our reactions to the activities. For instance when you count blades of grass, to use Rawls's famous example, you experience a sensation of boredom that then explains why you do not want to go on doing it.<sup>40</sup> On the Aristotelian view, pain and pleasure are *constituted* by our reactions towards our activities, including the activity of having certain experiences. They do not *explain* those reactions at all. On my view, this is all to the good, because in fact pleasure and pain do not play any essential role in the explanation of why we like or dislike the things that we do. What explains our reluctance to engage in the kind of activity that we call 'boring' is just the fact that it is too familiar, or too easy, or too monotonous, to engage the faculties. There is no need to say that those factors cause a painful sensation of boringness, which then in turn explains why you do not want to engage in the activity; the appeal to the sensation is completely otiose. The familiarity, easiness, difficulty, or whatever it is, causes the reaction directly. 'Boringness' is our name for one way of finding an activity painful, and finding it painful and wanting to stop doing it are pretty much the same thing. In some cases we cannot even say anything as vaguely explanatory as 'too easy' or 'too familiar' or 'too hard'. We are simply doing or undergoing something that we are wired not to want to do or undergo. At most, you can try to get your interlocutor to recognize the *character* of your objection: 'they taste kind of slimy to me'. But that is not the same as explaining it.

One thing that trips us up and gives the Benthamite view some specious plausibility is the fact that among the objects that we find painful or pleasant are certain sensations.<sup>41</sup> Sensations *are* among the objects of our experience: to take the philosopher's favourite, one can experience toothache. But on the Aristotelian view, the painfulness of that sensation is not the same as the sensation itself, nor is it an object of experience, at least in the first instance. The painfulness of the sensation rests in the fact that everything in you is working to flee from an object, in this case a sensation, that you cannot possibly flee from, since it emanates from your own nervous system. A person having painful sensations is trying to do battle with her own

<sup>40</sup> Rawls uses the example of someone whose only *enjoyment* rests in counting blades of grass, but the enjoyment of the activity is supposed to be unintelligible to most of us. See *A Theory of Justice*, pp. 379–380.

<sup>41</sup> Another is the fact that we can, of course, focus our minds on our pleasant and painful states, and in that sense can make them objects of experience. But then the object of experience is the way you are relating to your own condition. It is not something that can be detached from you and added on to the goods and bads of others the way Benthamite pleasures and pains can be.

shadow. She is trying to tear herself free from a condition that she cannot escape without the loss of her own self, precisely because it is a condition of herself. Pleasure and pain are not the objects of experience, but rather are themselves a *form* of experience—they are the way we experience our own condition, which is fundamentally evaluatively. They are, as Aristotle says in *On the Soul*, perceptions of what is good and bad.<sup>42</sup> In particular, they are perceptions of the goodness and badness of our evaluative condition.

## 6. WHY THERE IS SUCH A THING AS GOOD-FOR

This puts me in a position to explain my view. I began the last section by saying that there is such a thing as good and bad because there are such things as conscious animals. That is a hypothesis I share with the hedonistic utilitarians. But hedonistic utilitarians think this is because states of consciousness *themselves* are what can be good or bad—that they are the locus of value. I think it is instead because of the way in which conscious animals relate to their own condition—because of a special relation in which conscious animals stand to themselves, and in particular, to their own well-functioning, or their own evaluative goodness.

Earlier I pointed out that when we attribute health and other forms of fitness to an organism, we are describing a form of evaluative goodness. The organism which is fit and healthy will be good-for or good-at survival and reproduction, in an ordinary functional sense. But as a conscious organism, an animal stands in a special relationship to her own fitness. If she is to survive at all, she must have evolved in such a way as to monitor the relation between her own condition and what is going on in the world around her. She must, however roughly and defeasibly, respond positively to the things that promote her fitness and negatively to the things that threaten it.<sup>43</sup> It is only to the extent that animals do this that perception can contribute to fitness. The fact that the animal stands in this relation to her own relation to the world, to her fitness for it, is what explains the existence of the final good. In responding positively to the things that promote her own evaluative goodness, and negatively to those that harm it, an animal in effect *identifies* with her evaluative goodness, and this is what makes it *her* good. The things that are good-for her in the evaluative, functional sense present

<sup>42</sup> Aristotle, *On the Soul*, 3.7, 431a7–16.

<sup>43</sup> I am using Darwinian language here, but you do not have to live in Darwin's world in order to see this. The theory of pleasure and pain that I am advancing is similar to the one Hobbes advocates in *Human Nature*—roughly, that pleasure is the sense of something as empowering you and pain is the sense of something as enfeebling or undermining your power.

themselves to her as things that are worthy of pursuit; the ones that are bad as worthy of avoidance. Her awareness of her condition is an evaluative awareness, presenting her condition itself to her as something worth attending to. Here, as elsewhere, a form of consciousness or perception creates its own object.<sup>44</sup> That object is the final good.

Let me make it clear that I am not advancing what is sometimes called an evolutionary account of the good. I am not positing that what is good-for human beings, or for any other animals, is, in fact, whatever makes them healthy and fit in the evolutionary sense. Rather, my claim is that the final good is grounded in a relation in which conscious animals stand to their own evaluative goodness: they are motivated to monitor and attend to it, and, in that sense, to make it their end.<sup>45</sup> The other animals have little choice but to accept the purposes that nature sets for them, which is why I said earlier that they have no choice but to identify with the role that nature gives them. But that is not the case with us human beings. In the human case, the focus will switch to your fitness for carrying out the roles that I have elsewhere called your ‘practical identities’—the roles and relationships that give our lives meaning and value. Above all, the focus will switch to your fitness for constituting yourself as a unified human agent—that last being, according to my own view—the human role or function.<sup>46</sup> My point

<sup>44</sup> I mean that evaluative consciousness creates its own object in the same way that the senses create colours and sounds, for instance.

<sup>45</sup> A reviewer objects that all living things make their own evaluative goodness their end, since all living things adapt to external conditions in such a way as to maintain themselves. I think there is a difference between an entity’s being *adapted* so as to maintain its own condition, and its being *motivated* so as to maintain it. Admittedly, the content of this motivation is not the animal’s own good—a non-human animal obviously does not conceive and consciously aim at his or her good, any more than a plant does. But an animal aims at things that are good for her, and is motivated to pursue them. A plant is not an agent and it does not *aim at* anything, in the motivational sense. So its relation to the purposiveness of its nature is different from an animal’s—it’s evaluative good does not become its own good in the same way. I discuss the relationship between having a good and consciousness more extensively in ‘The Origin of the Good and Our Animal Nature’.

<sup>46</sup> For practical identities, see Lecture Three in *The Sources of Normativity* and *Self-Constitution: Agency, Identity, and Integrity*, ch. 1; for self-constitution as the human role, see *Self-Constitution* generally. Notice that what I just said implies the truth of the Aristotelian thesis that the moral virtues form part of your good, at least if Aristotle is correct in thinking that the moral virtues are essential to human functioning.

Some readers may be struck by the implication that a person’s good consists, not in having certain particular practical identities, but rather in being fit for the enactment of your practical identities, whatever they are. But I think that is for the most part true: the good-for a human being is to perform successfully and well in whatever roles she takes on. Of course, it can be true that the maintenance of a certain form of practical identity is called for by the overall human function of self-constitution—as in the case when it would be a violation of your integrity to give up a certain role, because of moral commitments you have taken on.

is simply that an entity's evaluative good becomes a final good when an entity is placed in her own keeping. That—placing an entity in her own keeping—is what conscious agency does.<sup>47</sup>

I advanced three complaints about hedonistic utilitarianism; I will now try to clarify my own view by saying why they do not hold for it. The first was the familiar complaint that according to utilitarianism, only states of consciousness are good or bad, and most of us agree that this is unintuitive. Objective conditions can be good or bad for us. This error, I maintain, arises in turn from the utilitarian's characteristic error, thinking of pleasure and pain as objects rather than forms of consciousness. The objects of consciousness are external conditions in so far as they bear on an animal's own condition. They are experienced as welcome when they promote the animal's health and fitness, or the person's enactment of her practical identity; they are experienced as unwelcome when they undermine those things. It is objective conditions, not mere states of consciousness, which have those effects.

I also worried about the question how evaluative and final goodness are related. As I said early in the chapter, I think most philosophers assume that when we talk about final goodness, what is involved is some sort of evaluation of our ends, and that is the relation between these two kinds of goodness. I think instead that when we talk about final goodness, we are talking about a relation in which we stand to our evaluative goodness, and to the objective conditions that enhance or promote it. Being a conscious agent makes our evaluative goodness our final good as well. If Aristotle is right in thinking that virtue is a form of evaluative human goodness—that is, is something essential to our functioning—then he is right about the connection between human goodness and human good.

This brings me to the third worry, the one I said was suggested by Moore's argument—that we cannot get from the notion of the final good, conceived as a kind of intrinsic value, to a notion of good-for at all. On my view, the

<sup>47</sup> There are philosophers who think that it is at least conceivable that there could be a conscious being who is not an agent, like a conscious rock. Those who hold the hedonistic utilitarian conception of pleasure and pain may also think such a being could (as a matter of conceptual possibility) experience pleasure and pain. Both Hobbes and Aristotle, whom I have invoked in trying to explain my view, would have denied this. For them, pleasures and pains are something like the first beginnings of motivated movement. The physiological theories in which these claims are embedded (described by Aristotle in *On the Movement of Animals*, and by Hobbes in *Human Nature*) seem rather primitive to a modern audience, but I believe these views are right to suggest a necessary connection between consciousness and agency. The idea that consciousness could simply inhabit a non-agent is dualistic, and springs from the same set of confusions about the nature of conscious experience that I am trying to describe in this chapter. To be pleased by something is to be inclined to go for it, or to ensure its continuance, not as a causal outcome, but constitutively. The idea is not separable from that of agency.

final good is something essentially relational, because it exists in relation to the consciousness, or to the point of view, of a conscious being. But this is not because the final good is a state of consciousness, as hedonistic utilitarians believe. It is because having a final good *just is* standing in a relation to evaluative goodness that is made possible by consciousness. So having a final good is an essentially relational state. All final goodness is essentially goodness-for the being whose final good it is. The question, then, is how we get from the notion of good-for to the notion of good *simpliciter*, if we can get there at all. I now turn to that question.

## 7. TOWARDS A CONSTRUCTIVIST ACCOUNT OF THE GOOD

I begin from a fourth objection to the hedonistic utilitarian account. The objection is that even if the view explains why there are good and bad things in the world, it does not explain why we operate with the concepts of good and bad. It is a thought tediously familiar to philosophers writing in the wake of G. E. Moore that pleasure and pain are merely naturalistic concepts, while good and bad are normative ones, or at least have normative implications. Something's being good is supposed to be some sort of a reason for or consideration in favour of it. This leaves the hedonistic utilitarian needing two things: an account of why we use the normative concepts at all, and an account of why we apply them to pleasure and pain in particular.

Although there have been non-cognitivist utilitarians,<sup>48</sup> hedonism most naturally suggests what I call a 'realist' account of the concepts of good and bad, by which I mean an account according to which we use these concepts because we encounter the good and the bad in our experience.<sup>49</sup> The concepts are essentially descriptive ones that pick out a property we find in the world, one that we can use in explanations of some kind. Hedonism also suggests an intuitionistic account of how we know that pleasure and pain are the natural phenomena to which these normative concepts apply. When we encounter pain, for instance, we 'see' right away that it is bad, and that is why we object to it.

But there is something oddly cerebral about this account. One might put the point by saying that every animal objects to pain, without needing the intervention of a normative characterization of the pain as something that is 'bad'. The other animals do not need to characterize pain as 'bad' before

<sup>48</sup> See for instance J. J. C. Smart in *Utilitarianism: For and Against*, p. 4.

<sup>49</sup> See Korsgaard, 'Realism and Constructivism in Twentieth-Century Moral Philosophy', for this account of what realism amounts to, especially pp. 302–303.

they object to it, so why should we need to do that? Reasoning about the badness of pain just seems like a pale copy of what is already going on. But, as I have already explained, I think the problem is even worse than that makes it sound, since I think that to say that an animal finds a condition painful is not to *explain* why the animal objects to it, but simply to *say* that the animal immediately objects to it. So it is not only that characterizing the painful situation as bad is not needed to *explain* why we object to it. Saying that we find our situation painful is already saying that we object to it. So what does the normative concept add?

In my paper, ‘Realism and Constructivism in Twentieth-Century Moral Philosophy’, I argue that the differences among so-called ‘metaethical’ positions should be understood as differences about why we use normative concepts rather than about the truth value of sentences or propositions.<sup>50</sup> I claim that according to realism we use normative concepts to describe or categorize normative facts or entities that we encounter in experience, while according to constructivism we use normative concepts as schematic markers for problems that we have to solve. For example, in Rawls’s theory of justice, the concept of ‘justice’ is used to refer to the way we solve the problem of how to distribute the benefits and burdens of social cooperation. One might add that according to expressivism we use normative concepts to express our endorsement of systems of values or rules of rationality or something of that kind. Different metaethical views are then different views about why we use normative concepts.<sup>51</sup> Why, thinking of things in these terms, does the constructivist think we use the concepts of good and bad?

In general, I believe that we use the concept of the good to mark out the solution to the problem of what things to aim at—what our ends should be, and of course derivatively what means we should take to them as well. That is a problem for rational beings, because we do not automatically go for the things that we desire. Our self-consciousness—our awareness of our own mental states—introduces a certain reflective distance between ourselves and our desires, so that we both can, and must, decide whether to act on them or not. So we face a problem—whether to act as our desires bid us. How do we solve that problem? I have already suggested that when we say that something would be good-for X in the final sense of good, we are speaking from a point of view made possible by our empathy, or in Hume’s sense, sympathy, with X. We are doing that because we are saying something that is only intelligible when we look at the world from X’s point of view.

<sup>50</sup> See ‘Realism and Constructivism in Twentieth-Century Moral Philosophy’, especially pp. 302–303; 308–309; 321–324.

<sup>51</sup> See Allan Gibbard in *Thinking How to Live*: ‘There is no contrast to be drawn between ethical and natural properties. The contrast is between ethical and naturalistic concepts’ (p. 105).

Claims about an animal's final good are essentially relative to the animal's point of view, for we identify his good and bad with respect to the view that he must take of his own condition.<sup>52</sup> Now it might seem a little odd to say that when I judge that something is good or bad for *me* I am sympathizing with myself, but in fact I think that this is exactly the right thing to say.<sup>53</sup> For to say that something is good-for me is to describe something's relation to my condition as having normative implications, and that in turn is to *endorse* the view of myself that, simply as a conscious being—as a being who is in her own keeping—I necessarily take of my own condition. One might see the endorsement of that view as an act of sympathy with myself.

However that may be, that act of sympathy or endorsement—the conscious determination to act on the motives that arise from the view that I necessarily take of my own condition—is what first gives rise to reasons. When I decide to treat what I cannot but think of as good-for me as being worthy of pursuit, then I decide to treat what is good-for me as good absolutely. In Kant's language, I treat myself as an end in itself. But that brings a new problem in its wake. For in a similar way, I believe, when I sympathize with the view that *another* necessarily takes of *himself*, I commit myself to the normativity of *his* final good. That's a large claim, and I cannot undertake to defend it in this short chapter, of course; I have discussed it elsewhere.<sup>54</sup> But on the assumption that we do find such reasons for actions in the final goods of others, then our judgements that various things are good, worthy of pursuit, both for ourselves and for others sets us a problem. The problem is that of finding something, if only it is possible, that would be good-for all concerned.<sup>55</sup>

The fact that we are faced with this problem then opens up the way for a constructivist account of why we use the concept of good *simpliciter*, and not just of good-for this or that person or animal. And this is exactly the account that we find in Kant. In the *Critique of Practical Reason*, Kant says, 'What we are to call good must be an object of the faculty of desire in the

<sup>52</sup> Of course X might disagree with our particular judgement about what his good is, but that is not a problem. X's view of his own condition is defeasible, but only a creature who has an evaluative consciousness of his own condition has a final good to be right or wrong about.

<sup>53</sup> I do not have space to work through the idea in this chapter, but I think one could construct a non-aggregative account of what is in an agent's own interests, and therefore of what is prudent for her to do, by thinking about the ends that she could share with herself at various different stages of her life.

<sup>54</sup> I refer primarily to my arguments in favour of the 'publicity' of reason in Lecture 4 of *The Sources of Normativity* and in *Self-Constitution*, ch. 9, but also to the claims I make about the consequences of our claiming the status of an 'end in itself' for ourselves in 'Fellow Creatures: Kantian Ethics and Our Duties to Animals', pp. 94ff.; and in 'Interacting with Animals', §8, pp. 107–109.

<sup>55</sup> See the account of the nature of shared deliberation in *Self-Constitution*, §9.4.

judgment of every reasonable human being, and evil an object of aversion in the eyes of everyone.<sup>56</sup> Kant does not mean that we somehow know that some things are good and some bad, and that therefore every reasonable human being must agree that they are such. Rather, he uses the idea as a criterion for identifying, or better, for constructing, the good itself: he goes looking for the ends that we can share. ‘Good’, then, is the schematic name for the solution to the problem of shared ends. This, then, is my answer to the question why we operate with the concept of the good: because as rational creatures who are capable of seeing the world through the eyes of others, we are faced with the task of constructing a state of affairs that is, as far as possible, good-for us all.<sup>57</sup>

## References

- Aristotle. *The Nicomachean Ethics*. Translated by W. D. Ross, translation revised by J. O. Urmson, in *The Complete Works of Aristotle*, edited by Jonathan Barnes. Princeton: Princeton University Press, 1984; cited in the usual way, with the Bekker page, column, and line numbers found in the margins of most translations.
- . *On the Movement of Animals*. Translated by A. S. L. Farquharson, in *The Complete Works of Aristotle*, edited by Jonathan Barnes. Princeton: Princeton University Press, 1984; cited in the usual way, with the Bekker page, column, and line numbers found in the margins of most translations.
- . *On the Soul*. Translated by J. A. Smith, in *The Complete Works of Aristotle*, edited by Jonathan Barnes. Princeton: Princeton University Press, 1984.
- Butler, Joseph. *Five Sermons Preached at the Rolls Chapel and A Dissertation upon the Nature of Virtue*. Edited by Stephen Darwall. Indianapolis: Hackett, 1983.
- Foot, Philippa. *Natural Goodness*. Oxford: Clarendon Press, 2001.
- Geach, P. T. ‘Good and Evil’, *Analysis* 17(2) (December, 1956), pp. 33–42.
- Gibbard, Allan. *Thinking How to Live*. Cambridge, MA: Harvard University Press, 2003.
- Hobbes, Thomas. *Human Nature and De Corpore Politico*. Edited by J. C. A. Gaskin. Oxford: Oxford University Press, 1994.
- Hume, David. *A Treatise of Human Nature*. Second edition edited by L. A. Selby-Bigge and revised by P. H. Nidditch. Oxford: Clarendon Press, 1978.
- Kant, Immanuel. *Critique of Practical Reason* (Cambridge Texts in the History of Philosophy). Translated and edited by Mary Gregor with an Introduction by Andrews Reath. Cambridge: Cambridge University Press, 1997; cited in the usual

<sup>56</sup> Kant, *Critique of Practical Reason*, 5:61.

<sup>57</sup> I would like to thank audiences at the 8th Annual Metaethics Workshop and on the occasion of the Lindner Lecture at Wooster College for discussion, and two anonymous reviewers for comments. I would especially like to thank Melissa Barry and Connie Rosati, who were my commentators at Wooster.

- way with the Prussian Academy edition volume and page numbers found in the margins of most translations.
- Korsgaard, Christine M. 'Aristotle on Function and Virtue', in *The Constitution of Agency*. Oxford: Oxford University Press, 2008, 151–173.
- 'Aristotle's Function Argument', in *The Constitution of Agency*. Oxford: Oxford University Press, 2008, 129–150.
- 'Fellow Creatures: Kantian Ethics and Our Duties to Animals', in *The Tanner Lectures on Human Values*. Edited by Grethe B. Peterson. Salt Lake City: Utah University Press, vol. 25, 2005; and available at <[http://www.tannerlectures.utah.edu/lectures/documents/volume25/korsgaard\\_2005.pdf](http://www.tannerlectures.utah.edu/lectures/documents/volume25/korsgaard_2005.pdf)>.
- 'Interacting with Animals', in *The Oxford Handbook of Animal Ethics*. Edited by Tom Beauchamp and R. G. Frey. Oxford: Oxford University Press, 2011, 91–118.
- 'The Myth of Egoism', in *The Constitution of Agency*. Oxford: Oxford University Press, 2008, 69–99.
- 'On Having a Good', forthcoming, in *Philosophy: The Journal of the Royal Institute of Philosophy*.
- 'The Origin of the Good and Our Animal Nature', forthcoming, in *Problems of Goodness: New Essays on Metaethics*. Edited by Bastian Reichardt.
- 'Personal Identity and the Unity of Agency: A Kantian Response to Parfit', in *Creating the Kingdom of Ends*. New York: Cambridge University Press, 1996, 363–397.
- 'Realism and Constructivism in Twentieth-Century Moral Philosophy', in *The Constitution of Agency*. Oxford: Oxford University Press, 2008, 302–326.
- *Self-Constitution: Agency, Identity, and Integrity*. Oxford: Oxford University Press, 2009.
- *The Sources of Normativity*. Cambridge: Cambridge University Press, 1996.
- 'Two Distinctions in Goodness', in *Creating the Kingdom of Ends*. New York: Cambridge University Press, 1996, 249–274.
- Moore, G. E. *Ethics*. Oxford: Oxford University Press, 1912; paperback edition, 1965.
- *Principia Ethica*. Cambridge: Cambridge University Press, 1903.
- Nozick, Robert. *Anarchy, State, and Utopia*. New York: Basic Books, 1974.
- Parfit, Derek. *Reasons and Persons*. Oxford: Clarendon Press, 1984.
- Plato. *Republic*. Translated by G. M. A. Grube and revised by C. D. C. Reeve, in *Plato: Complete Works*. Edited by John M. Cooper. Indianapolis: Hackett, 1997; cited in the standard way with Stephanus numbers.
- Rawls, John. *A Theory of Justice*. Cambridge, MA: Harvard University Press, 1st edition, 1971; revised edition, 1999.
- Smart, J. J. C. and Bernard Williams. *Utilitarianism: For and Against*. Cambridge: Cambridge University Press, 1973.