Darwall, Stephen, *Welfare and Rational Care*, Princeton: Princeton University Press, 2003, pp. xi + 135, US$24.95 (cloth).

In this slim volume, Stephen Darwall advances three main theses on the topic of welfare, or prudential value—the topic of what makes a person's life go well for him or her. Darwall defends (i) a theory in the metaethics of welfare, which he calls the 'rational care theory of welfare'; (ii) the thesis that welfare has 'agent-neutral normativity' rather than 'agent-relative normativity'; and (iii) a partial normative theory of welfare, which he calls the 'Aristotelian Thesis'. In addition, Darwall discusses the psychological concept of care, which figures as a conceptual primitive in the rational care theory of welfare.

Most of the book has appeared previously. Chapter II is a slightly modified version of Darwall's 'Self-Interest and Self-Concern' [1997]. 'Empathy, Sympathy, Care' [1998] provides the content and the title of Chapter III. And 'Valuing Activity' [1999] appears, with minor modifications, as the fourth and final chapter. Only Chapter I is new.

I was disappointed with the book. Having admired some of Darwall's earlier work, and being especially interested in welfare, I looked forward to the new book. But I found the style often loose and obscure, many of the arguments weak or unoriginal, and many of the claims either implausible or, when plausible, either not very original or rather modest. Furthermore, the central positions are, in my view, underdeveloped—they stand in need of clarification, exploration, and/or extension. Below I focus on the Aristotelian Thesis and the rational care theory of welfare.

The Aristotelian Thesis: Darwall's 'Aristotelian Thesis' identifies something that is good for human beings, namely, engaging in activities in which one *appreciates* the *worth* or *merit* of things with worth or merit. Darwall says in passing [80] that this may not be the only fundamental source of human welfare. Since the Aristotelian Thesis (i) is about human welfare only, (ii) is a sufficient condition only, (iii) says nothing about what is bad for a subject, and (iv) contains no quantitative principles, it is far from a complete theory of welfare.

Worth and Merit. No full account of either worth or merit is given, but a few sufficient conditions and some examples shed some light on the ideas. A thing has worth if it has moral considerability (so all persons have worth, according to Darwall [78–9, 83]). A thing has worth if it is beautiful [79, 102]. A thing has *merit* if (and only if, I think) 'it is appropriately related to worth' [89]. So acts of creating or appreciating a beautiful thing have merit because creation and appreciation are appropriate relations to bear to beautiful things, and beautiful things have worth [79, 102]. For analogous reasons, acts and attitudes of respect and caring for a person have merit [79, 94].

Appreciation. To appreciate the worth or merit of something is more than merely to believe truly that it has worth or merit [90, 95–6]. Appreciation consists, at least in part, in feelings and emotions [84–5, 90]. And it is an evaluative attitude, an 'evaluative response' to a thing of value [89]. In an insightful passage [84–5], Darwall distinguishes the pleasurable feeling of a warm shower on a cold day from the pleasurable feeling of pride. Though both feelings may be good in themselves, only the latter case is a case of appreciation, since only the latter entails seeing something else (in this case, whatever the person is proud of) as good.

Despite its name, the Aristotelian Thesis bears a more obvious resemblance to Moore's ideas about the goodness of the appreciation of beautiful objects and of admirable people [Moore 1929: 188–200].[1] One is also reminded of Parfit's 'composite' theory that welfare consists in wanting and getting (and/or enjoying) the things on the 'objective list' [Parfit 1984: 501–2].[2] Darwall acknowledges the connection to Moore and to Parfit [102, 120n], but seems to want to distance himself from views with a hedonistic flavour. Whereas for Moore, aesthetic appreciation includes enjoyment as a *component*, Darwall suggests [76, 95] that the relation between appreciation and pleasure is merely *causal*. This distance from hedonism opens even the modest Aristotelian Thesis up to an objection, an objection that confronts any theory of welfare that fails to include pleasure or enjoyment as an essential element. The objection is that, according to such theories, there are possible lives containing no enjoyment—and even, perhaps, constant boredom—that are nevertheless great lives for the people who live them. The dialectical advantage of hedonism is observed by Robert Adams [1998: 95]:

> Another truth about human well-being that is intuitively evident is that a person's good is not very fully realized unless she likes or enjoys her life in the long run. You may be very virtuous; you may be brilliant, beautiful, successful, rich, and famous, but if you do not enjoy your life, it cannot plausibly be called a good life *for you*. We may think of this as the kernel of truth in hedonism. . . .

The Rational Care Theory of Welfare: The rational care theory of welfare is an analysis of the concept of welfare: it alleges to provide the meaning of sentences of the form '*x* is good for *S*'. Darwall puts the theory in quite a few different ways [4, 7, 8, 8–9, 12, 31, 45, 46, 48, 53, 71, 83]. One statement of it reads, 'what it is for something to be good for someone *just is* for it to be something one should desire for him for his sake, that is, insofar as one cares for him' [8]. I formulate the theory as follows:

> RCTW: *p* is good for *S* = df. for any person *x*, if *x* cares for *S*, then *x* should desire that *p* be true for *S*'s sake.[3]

The theory thus reduces the concept of welfare to the concepts of care, desire, and the concept expressed by the 'should', which, Darwall says, is to be taken in 'its most general normative sense' [8]. Darwall implies that '*x* should want' here would mean the same as 'it makes sense for *x* to want' [45], 'it is rational for *x* to want' [9], and '*x* has reason to want' [71]. Darwall suggests that only the concept of care is primitive in RCTW [71], but it seems to me that the 'should' is primitive here as well—I can find no account of it anywhere in the book.

Two objections to RCTW that come immediately to mind are anticipated by Darwall, but I find his responses unsatisfactory. I call one 'the circularity objection' and the other 'Darwall's Euthyphro problem'.

The Circularity Objection. The first objection is that RCTW is covertly circular, because surely any plausible definition of 'care' will make use of the concept of welfare, a concept in the analysandum of RCTW. Darwall acknowledges that a definition of 'care' in terms of welfare is unavailable to him. He has no alternative definition to give, and so takes, or seems to take, the following strategy: he attempts to establish that care is a natural kind (this is the stated goal of Chapter III); and this (it is implied) relieves him of the duty to define 'care'. Though Darwall never says it, perhaps his idea is that the word 'care', since it refers to a natural kind, functions like a tag or a proper name and therefore has no meaning. And though Darwall never says it, perhaps he thinks we (or the experts in psychology) can identify the psychological state of care because it figures in psychological laws. Instead of explaining

---

[1]   But wasn't Moore sceptical about the whole idea of welfare, of something being good *for someone*? No—he merely denied that welfare is a kind of value distinct from and irreducible to value *simpliciter* [Moore 1929: 97–9].

[2]   Parfit outlines such a theory, and speaks favourably of it, but does not commit himself to it. Other recent views along the Moorean and Parfitian lines include Robert Adams's theory that welfare consists in enjoyment of the excellent [1999: 93–101] and Fred Feldman's theory that welfare consists in taking pleasure in the pleasure-worthy [Feldman 2002: 604–27].

[3]   Darwall almost invariably uses the expression 'insofar as' in stating the theory ('a person's good is constituted . . . by what one (perhaps she) should want *insofar as one cares about her*' (4)), but its meaning is never made clear. Perhaps Darwall is saying that care and desire come in degrees, and that, if g is good for S, then if you care for S to degree n, then the degree to which you want g for S's sake should be n. I will stick with the interpretation on display above since it is weaker (i.e., it is entailed by the more complicated interpretation) and since I am more confident that Darwall would accept it than I am that he would accept the more complicated interpretation.

My formulation presupposes that the bearers of welfare are propositions. It is not clear to me whether Darwall would accept this, but nothing that follows hangs on it.

why I find Darwall's execution of this strategy (if this is indeed his strategy) unconvincing, I want to pose a dilemma.

Not only must Darwall reject a definition of 'care' in terms of welfare, he cannot even accept the following thesis as analytic:

CW: necessarily, if *x* cares for *y* at *t*, then *x* is disposed to desire at *t* that *y* fares well.

It *is* open to Darwall to accept a weakened version of CW, one whose modal status is reduced to mere causal necessity. So perhaps Darwall would claim that, as a matter of contingent psychological fact, whenever anyone cares for another, he is disposed to desire what is good for her. (Darwall does accept *some* sort of relationship between care and welfare, though he leaves it very vague, as when he says that care 'is a feeling or emotion that...*involves* concern for [an individual], and thus for his welfare, for his sake.' [51, emphasis added].)

In any case, to claim that the connection between care and a desire for welfare is merely contingent (i.e., to deny that CW is conceptually necessary) has an implication that is difficult to swallow. The implication is that it is conceivable for someone to care deeply for another person—really to care for him—while nevertheless having no desires whatsoever—not even dispositional desires—that the person fare well, that the person be ok, that the person not be harmed. This is like trying to conceive of a married bachelor.

Darwall is thus faced with a dilemma. Either care is analytically tied to welfare (in the manner of CW), or it is not. If it is, RCTW would seem to be circular, and in a vicious way. If it is not, then care without the slightest interest in welfare is a possibility.

Darwall's Euthyphro Problem. The second objection is that RCTW gets the order of explanation backwards: things aren't good in virtue of the fact that we should want them; rather, we should want them because they are good. Suppose state of affairs *g* is good for *S*. Perhaps we can all agree that, in typical cases at least, if I care for *S*, then I should want *g* to obtain for *S*'s sake. But why should I want *g* to obtain? The natural answer, of course, is that I should want *g* to obtain because *g* is good for *S*. My complete reason for wanting *g* to obtain is: that *g* is good for *S*, that I care about *S*, and that to care about someone is (at least in part) to desire that he fares well (i.e., gets what is good for him). This natural answer, of course, is unavailable to Darwall.

We could put the problem socratically. Should we (who care about *S*) want *g* to obtain because *g* is good for *S*? Or is *g* good for *S* because we should want *g* to obtain? The fact that RCTW takes the latter, less intuitive, horn is Darwall's Euthyphro Problem.

Darwall acknowledges his Euthyphro Problem but argues against taking the more intuitive horn [9–11]. Upon analysis, the argument appears to rest covertly on the following principle: no concept that is 'normative for' some other concept is 'itself an explicitly normative concept'. But why think this is true? *Goodness of outcome* seems to be both 'normative for' some other concept while being 'itself an explicitly normative concept'. It is 'normative for' action—i.e., that some outcome would be good provides a reason (albeit an overridable reason) to perform the action that would bring about that outcome. And how good an outcome is, is surely 'itself an explicitly normative concept', as Darwall understands it.[4]

I think there are other problems with the book. There are the strange and, in my view, unsuccessful arguments for the claim that the Aristotelian Thesis and the rational care theory of welfare 'mutually support' each other. One could also complain that the meaning of the thesis that welfare has 'agent-neutral normativity' is never made clear. These charges will have to go unsupported here.[5]

Christopher Heathwood   *University of Massachusetts, Amherst*

REFERENCES

Adams, Robert Merrihew 1999. *Finite and Infinite Goods: A Framework for Ethics*, New York: Oxford University Press.

Darwall, Stephen 1997. Self-Interest and Self-Concern, *Social Philosophy and Policy* 14: 158–78.

[4]   Which seems to be as follows: a concept is itself an explicitly normative concept iff two people who agree about all of the non-normative facts can coherently disagree about whether or how that concept is instantiated [11].

[5]   I thank Fred Feldman, Kristen Hine, Jens Johansson, Kris McDaniel, Jason Raibley, and Michael Rubin for discussing the book with me. Special thanks to Fred Feldman, Kris McDaniel, Jason Raibley, and Stephan Torre for helpful comments on an earlier draft.

Darwall, Stephen 1998. Empathy, Sympathy, Care, *Philosophical Studies* 89: 261–82.

Darwall, Stephen 1999. Valuing Activity, *Social Philosophy and Policy* 16: 176–96.

Feldman, Fred 2002. The Good Life: A Defense of Attitudinal Hedonism, *Philosophy and Phenomenological Research* 65: 604–27.

Moore, G. E. 1929. *Principia Ethica*, Cambridge: Cambridge University Press.

Parfit, Derek 1984. *Reasons and Persons*, Oxford: Clarendon Press.