

J. DAVID VELLEMAN
Is Motivation Internal to Value?*

Abstract: The view that something's being good for a person depends on his capacity to care about it – sometimes called internalism about a person's good – is here derived from the principle that 'ought' implies 'can'. In the course of this derivation, the limits of internalism are discussed, and a distinction is drawn between two senses of the phrase "a person's good".

1. Introduction
 2. The Intuitive Basis of Internalism
 3. What the Intuitive Basis Will Support
 4. The Concept of What Can Matter to Someone
 5. Mattering vs. Motivating
 6. Why We Need Two Concepts of a Person's Good
 7. Two Concepts of the Good, Two Versions of Internalism
 8. Internalism as an Instance of 'Ought' Implying 'Can'
 9. 'Prima Facie Ought' and 'Prima Facie Can'
 10. What This Derivation Implies
 11. Conclusion
- References

1. *Introduction*

Various philosophers have thought that the truth-conditions of some normative statements – the statement that something is good for a person, or that an action is morally required of him, or that there is some reason for him to act – include or entail that the person would be moved to act accordingly under some conditions.¹ I want to argue that there are good reasons for accepting one version of internalism, as this thesis is usually called; but that if these reasons are all that can be said for internalism, as I suspect, then its scope and implications are more limited than they are often conceived to be.

The version of internalism that I'll try to vindicate is restricted to statements about what's intrinsically good for a person. One reason for focusing on intrinsic value for a person is that it is, of all the normative properties, the

* For comments and discussion, I am grateful to the participants in the *Preferences* conference, and especially to my commentator, Georg Meggle.

¹ For general discussions of this view, see Falk (1986), Frankena (1958), Darwall (1983), pp. 51–61 et passim.

one of which internalism seems most likely to be true. The norms of morality and rationality aren't tailored to suit individual tempers; and so we can imagine an agent expressing indifference to morality, and perhaps even to rationality, by saying that they aren't for him. But a person's own good is indeed tailored to him, and we cannot imagine him saying that his good is not for him, since its being for him seems essential to its being specifically his good.² As Peter Railton has said, "it would be an intolerably alienated conception of someone's good to imagine that it might fail in any [...] way to engage him".³ My strategy, then, will be to examine what truth there is in internalism about a person's good, on the assumption that there will be some truth in that version of internalism if there is any truth in the doctrine at all.

2. *The Intuitive Basis of Internalism*

I think that Railton puts his finger on the intuitive basis of internalism about the good in the following passage:⁴

"[N]otions like good and bad have a place in the scheme of things only in virtue of facts about what matters, or could matter, to beings for whom it is possible that something matter. Good and bad would have no place within a universe consisting only of stones, for nothing matters to stones. Introduce some people, and you will have introduced the possibility of value as well. It *will* matter to people how things go in their rock-strewn world."

These reflections imply that the existence of value entails the existence of beings to whom things can matter. And this conclusion sounds like a version of internalism about the good.

Railton's intuition is similar to one that other philosophers have adduced in support of very different conclusions. Sidgwick, for one, claimed that things "are not ultimately and intrinsically desirable [...] when considered apart from any relation to conscious existence" – a claim that sounds very much like Railton's.⁵ "For example", Sidgwick said, "we commonly judge

² For the idea that a person's good is tailored to him, see Rosati (1995).

³ Railton (1986), p. 9. Note, however, that my ellipsis materially affects Railton's meaning. What Railton says is that it would be an alienated conception of someone's good "to think that it might fail in any *such* way to engage him," the way in question being "if he were rational and aware" (my emphasis). I think that a person's good might still fail to engage him under these particular conditions – for example, if he were himself suffering alienation, which is perfectly compatible with rationality and awareness. I discuss these issues at greater length below.

⁴ *Ibid.*

⁵ Sidgwick (1907), p. 401.

some inanimate objects [...] to be good as possessing beauty, and others bad from ugliness: still no one would consider it rational to aim at the production of beauty in external nature, apart from any possible contemplation of it by human beings".⁶ One can almost hear Sidgwick adding that aesthetic value would have no place in a universe consisting merely of stones.

Unlike Railton, however, Sidgwick employed this intuition in the service of hedonism. From the premise that things have intrinsic value only "in relation to conscious existence" he concluded that the requisite relation is that of identity, in the sense that states of consciousness are the only things capable of having value. And this conclusion seems more than the intuition by itself (or perhaps any reasonable intuition) can support.

In light of the uses to which the intuition expressed by Railton has been put, then, we would be wise to consider carefully what exactly it implies. If Railton has accurately identified the intuitive basis of internalism, the question remains what exactly that basis will support. What are we entitled to conclude from the intuition that a world of stones would be devoid of value?

3. *What the Intuitive Basis Will Support*

One might claim that we're entitled to draw no more than the weakest conclusion, to the effect that nothing can be good or bad unless there is someone to whom something can matter. This conclusion would not require that the power of mattering to a creature be vested in the object that has value; nor would it stipulate that an object can have value only for those creatures to whom it can matter. So long as something can matter to someone, according to this conclusion, anything can be good or bad for anyone, or valuable absolutely.

Yet this conclusion is surely less than Railton's intuition calls for. Railton's story of the unpopulated world suggests, to begin with, that introducing potential subjects of concern into a world lays a basis for value only in the potential objects of their concern, since objects that can't arouse their concern remain, as before, in an affective vacuum. Things that cannot matter to the only creatures to whom anything can matter might as well be back among the stones.

The story also suggests that introducing potential subjects of concern into a world lays a basis for value only by introducing potential subjects of value – that is, creatures for whom things can be good or bad. The arrival of sentient

⁶ *Ibid.*, p. 114.

beings in a world of stones wouldn't render anything potentially good or bad *for* the stones, since the stones would remain as impassive as before. And if things could have value absolutely – a value that didn't consist in being good or bad for someone – then why would that value depend on their chances of mattering to sentient beings? Absolute value is precisely the sort of value that something ought to possess even if it were the only thing in the world. If having value requires bearing a relation to the potential concerns of sentient beings, then value would appear to be essentially relational.

Thus, Railton's story suggests the stronger conclusion that nothing is good or bad unless it can matter to someone, and that it is then good or bad specifically for him. The story suggests, in other words, that a valuable object must itself be a potential object of concern and has value only for the potential subjects of that concern.

Now, I am not currently interested in whether value can be absolute or must instead be relative to a subject. The question that interests me is whether value entails a relation to anyone's potential concerns. And if any kind of value is likely to entail such a relation, as I've already suggested, it's the kind that's relative to the subject in question. That is, the kind of goodness that's most likely to depend on whether the object possessing it can matter to a creature is the quality of being good for that creature.

4. *The Concept of What Can Matter to Someone*

The conclusion that nothing can be good for someone unless it can matter to him is still relatively weak, in various respects. For instance, it does not say that in order for something to be good for someone, he must be disposed to care about it under particular conditions that are specifiable without reference to the good. All it says is that his caring about it must be a possibility.

Furthermore, Railton's story suggests that the requisite possibility may be relatively remote. What distinguishes stones from people is that stones are necessarily devoid of concern: nothing capable of caring about anything would qualify as a stone. All that Railton's story implies, then, may be that something can't be good for someone if he's constitutionally incapable of caring about it – if caring about it is beyond the affective capacities of a creature like him. And this conclusion allows that the thing can be good for him so long as the possibility of his caring about it isn't ruled out by his very nature.

The resulting conclusion doesn't entirely lack significant consequences. Some sentient creatures are constitutionally incapable of caring about some things, because those things are necessarily beyond their grasp. For example,

a cat cannot care about being famous; nothing capable of caring about fame would qualify as a cat (rather than as a person in feline form). The present conclusion therefore implies that fame is not intrinsically good for a cat – a consequence that might bear significantly, say, on the humane treatment of show animals.⁷

Unfortunately, we cannot similarly specify the analogous consequences that this conclusion might yield for people. If some things lack human value because they lie beyond human grasp, we humans are in no position to say what they are. There might of course be things that lie within our comprehension but beyond the possible scope of our concern; and the present conclusion would be a premise from which to prove that such things cannot be good or bad for us. But one is hard pressed to think of interesting examples.

5. *Mattering vs. Motivating*

The present conclusion is also weak in that it speaks of what matters to a creature rather than what the creature wants or is moved to pursue. Hence the version of internalism suggested by Railton's story doesn't posit a connection between having value and motivating; it posits a connection between having value and mattering.⁸

The difference is that what matters to a creature includes things that have already come about, whereas desire and motivation are ordinarily restricted to things that haven't. You cannot want something to be the case, or be moved to make it the case, if you know that it already is the case. But you can still be glad that it is the case; and if you're glad about it, then it matters to you, even if you cannot retrospectively desire it and couldn't have desired it in advance.

Suppose there were a kind of experience that you would have to imagine in order to want, and that you would have to undergo at least once in order to imagine. You might then be constitutionally incapable of desiring your very first experience of this kind. For until you'd had the first such experience, you'd be unable to want it yet, because you couldn't imagine it; and once you'd had the first one, you'd be unable to want that one any more, since you would unmistakably have had it.⁹ Would we say that your first experience of this kind couldn't be good for you, because you were incapable of wanting it? Surely,

⁷ For an argument of this form, see my (1991).

⁸ On the difference between mattering and motivating, see Elizabeth Anderson (1993), ch. 2.

⁹ See Griffin (1986), p. 11: "I might get something I find that I like but did not want before because I did not know about it, nor in a sense want now simply because I already have it". See also Hare (1989).

your capacity to care about the experience retrospectively, by being glad to have had it, should satisfy any reasonable constraint on what can have value for you.

I am therefore tempted to reject the formulation of internalism as a thesis about the relation between value and motivation. I am inclined to formulate it instead as a thesis about the relation between value and affect, which encompasses motivation and more.

6. *Why We Need Two Concepts of a Person's Good*

The distinction between interpreting internalism as a thesis about motivation and interpreting it as a thesis about affect corresponds, I think, to a distinction between senses of the phrase "a person's good". In one sense, this phrase refers to whatever constitutes a person's welfare or well-being. But philosophers sometimes use the phrase in a different sense, with very different implications.

Consider, for example, Sidgwick's claim that "a man's future good on the whole is what he would now desire and seek on the whole if all the consequences of all the different lines of conduct open to him were accurately foreseen and adequately realised in imagination at the present point of time"¹⁰. Some qualifications must be read into this passage, since Sidgwick has previously limited his discussion to "what a man desires for itself – not as a means to an ulterior result – and for himself – not benevolently for others"¹¹. Yet even so qualified, Sidgwick's account cannot mean that whether a particular future outcome will be good for a person depends on whether he would want and work toward that outcome at present. For in that case, the account would attach different and incompatible values to future outcomes as changes occurred in a person's present dispositions to desire and pursue them. You can be alternately attracted and repelled by the prospect of some future event; but that event cannot be both your future good and your future ill, in a sense that

¹⁰ Sidgwick (1907), pp. 111 f.

¹¹ *Ibid.*, p. 109. Here Sidgwick implies that the structure of desires is more complex than is sometimes acknowledged. All philosophers recognize the possibility of wanting one thing for the sake of something else that one wants. In this case, the former object is desired extrinsically, as a means, and one's desire for it is psychologically dependent on one's desire for the end. What Sidgwick suggests, however, is the possibility of wanting something intrinsically but for the sake of a person. In this case, the phrase "for the sake of" does not introduce a relation of dependence between two desires whose objects are regarded as end and means. Rather, it introduces, so to speak, an indirect object of the desire, in the form of a person for whom something is wanted.

would entail its being both good and bad for you at the time of its future occurrence.

Of course, Sidgwick's account would identify your future good with what you would desire only after being informed about the motivational changes that you were due to undergo (among other matters); but it would still make your future good depend on how that information would impress you as you are now. Suppose that you were once disposed to desire and pursue some future state of affairs even in light of the knowledge that you were destined to lose that disposition, and even though you did in fact lose it. In that case, the state in question was your future good at one time, according to Sidgwick's account, but then ceased to be your future good. Yet if that state of affairs comes about, it will then either be or not be a state of well-being for you, in what was at both times your future. Hence what would constitute someone's welfare in the future cannot be what Sidgwick means by the phrase "a man's future good"¹².

What's more, the things that a person is currently disposed to desire and seek may not exhaust the constituents of his welfare, because his welfare may include things whose value depends on their having been unanticipated and unsought.¹³ If there are windfalls that would be good for a person precisely in virtue of being windfalls – things such as unsolicited affection or spontaneous merriment – then no amount of reflection on the consequences of potential actions would lead the person to desire or seek them, since the constituents of these goods would lose their attraction when considered as consequences of his own efforts.¹⁴

¹² The problem I have raised is cited in Brandt (1982) as a reason for rejecting desire-based conceptions of utility or well-being. I prefer to regard the problem as an indication that there is more than one sense of the phrase "a person's good". In Brandt (1989), p. 40, he suggests that conceptions of utility based on fully informed desires may elude this problem, because the fully informed desires that a person is disposed to have at different times would in fact tend to converge. But this empirical claim is inadequate to solve the problem. For we do not conceive of well-being as something whose determinateness depends on some contingent fact about a convergence among one's past dispositions to desire.

¹³ See Griffin (1986), p. 22: "Good things can just happen; manna from heaven counts too." I am making the somewhat stronger suggestion that the value of things may sometimes depend on the very fact that they "just happen" – on their being not only manna but manna *from heaven*. See Elster (1983), pp. 44–52. The phenomenon at issue here is related to what Parfit (1984) calls "self-effacingness".

¹⁴ Immediately before the passage that I first quoted, Sidgwick considers a slightly different objection: "[A] prudent man is accustomed to suppress, with more or less success, desires for what he regards as out of his power to attain by voluntary action – as fine weather, perfect health, great wealth or fame, etc.; but any success he may have in diminishing the actual intensity of such desire has no effect in leading him to judge the objects desired less 'good'" (1907, p. 110). Sidgwick therefore modifies his initial account by identifying a person's good

The latter problem disappears, however, if Sidgwick's phrase "a man's good" is interpreted to mean "the proper object of self-interest" – that is, the goal that would be rational for a person to aim at for his own sake, insofar as he aims at anything for himself.¹⁵ What a person would pursue in a spirit of self-interest under conditions of full information may well be all that he rationally ought to pursue in that spirit, even if other things might make for his well-being when acquired without effort.

The proposed interpretation may also remove the former problem, since it renders "a person's future good" as "the proper future object of a person's self-interest" – meaning, that future goal which would be rational for him to pursue for himself. Although one and the same future state cannot both be and not be a state of well-being for a person at the time of its occurrence, it can still be rational for him to pursue at one time and not at another. For there can be (prudential) reasons for pursuing something self-interestedly other than the fact that one would benefit from the thing's attainment. The self-interested pursuit of something may be intrinsically valuable for a person, since the pursuit itself may be one component of a good life; or it may have the extrinsic value of providing incidental satisfactions and opportunities for growth.

Consequently, a person sometimes regards the attainment of his former goals as not truly in his interest without regretting that he pursued them, and without wishing that he had always preferred the goals that he now prefers. Indeed, a person may even be glad that he didn't make a youthful beeline for what he has now resolved to pursue self-interestedly; and he may be quite confident that his pursuing it now is rational, even though he's uncertain whether its future attainment will serve his interests later on.

Thus, for example, I came to my career in philosophy circuitously, via several years of preparation for a career in a different field. I now believe that being a classicist in my thirties – which is what I originally aimed at – would not have been at all good for me. Yet I'm glad that I aimed at it and that I didn't

with what he would want "if it were judged attainable by voluntary action" (p. 111). But for reasons that I don't fully understand, this seemingly essential qualification drops out in Sidgwick's final formulation of the account, which specifies only that the agent is to consider the possible consequences of actions that are actually open to him. In any case, the qualification in question wouldn't solve the problem that I have raised in the text.

¹⁵ Here I am assuming, with Sidgwick, that we can speak of *desiring* and *pursuing* something 'for oneself', and I am referring to such attitudes as constituting one's self-interest. See note 11, above.

Note also that the terms "ought" and "rational" are here being used in their prudential senses. Hence when I speak of what a person ought to pursue self-interestedly, or what would be rational for him to pursue self-interestedly, I mean what the norms of rational self-interest would recommend that he self-interestedly pursue.

initially set out to be a philosopher instead. For I think that my actual life, in which I set out to become a classicist and then became a philosopher, has turned out to be better than either sticking with the Classics or starting out in Philosophy would have been. And I think so largely because I think that seeking to be a classicist was the better way for me to spend my twenties – and that the person who wanted to be a classicist was the better person for me to be in my twenties – even though being a classicist wouldn't have been a good way to spend my thirties. I therefore think that being a classicist now, though not in my interest now, was the proper future object of my self-interest fifteen years ago and was therefore, in a sense, my future good at the time. In this sense, my future good can change from one time to another, even though I can have only one future and only one future set of interests.

7. *Two Concepts of the Good, Two Versions of Internalism*

Reflection on Sidgwick's account of a person's future good thus reveals that we speak of a person's good in two subtly different senses, corresponding to two different perspectives in which personal evaluations take place. On the one hand, a person deliberates about what to aim at for his own sake, and we consider how to advise him in such deliberations or how to judge his success in them. From this perspective, the person's good is whatever he ought to seek for himself – what I have called the proper object of his self-interest – including things whose pursuit is valuable independently of their attainment, but not things whose value depends on their not being pursued.

On the other hand, we sometimes evaluate how well someone is doing, or how a particular action on our part would impinge on his welfare. From that perspective, our concern is not to guide the person's own self-interested motives to their proper object; it is simply to evaluate how the person is faring or would fare under various possible outcomes. Things that the person rationally ought to seek for his own sake might not benefit him by their occurrence; whereas he might benefit from the occurrence of things that he has no reason to seek. Hence the person's good appears slightly differently from this latter perspective.

These two senses of "a person's good" might be labeled the agent's and patient's senses, since one is grounded in the perspective of practical reasoning, and the other, in the perspective of the person as he is affected by events. But since a person's good in the patient's sense is also what we ordinarily call his well-being, I shall mark the distinction by referring to a person's practical good, on the one hand, and his well-being, on the other.

Now, if we plan to apply internalism to a person's practical good, we may want to formulate it as a thesis about the relation of value to motivation rather than affect in general. A person's practical good, after all, is the proper object of his self-interested motives, and we might therefore expect that it can exist only in relation to such motives, whatever the relevant relation may be. Since a person's well-being is not necessarily the proper object of any motives, it should require no relation to motives.

Yet a person's well-being is the proper object of his self-regarding affect in some sense, since it is intrinsically valuable for him and hence something he ought to value for himself – something about which he ought to be at least retrospectively glad, for his own sake, though not necessarily prospectively desirable. Perhaps, then, a person's well-being depends, not on a relation to his self-interested motives, but rather on a relation to his self-interested affect in general.

8. *Internalism as an Instance of 'Ought' Implying 'Can'*

This thought can be formulated more clearly as follows. Say that a person's practical good is that which ought to be the object of his self-interested motives – what he ought to desire and be inclined to pursue for his own sake; whereas his well-being is that which ought to be the object, more generally, of his self-regarding affect, or self-concern – what he ought to cherish or treasure or be glad about, in the same, self-regarding spirit. Naturally, what *ought* to be the object of an attitude must be something that *can* be the object of that attitude: it cannot be the case that a person ought to have some attitude toward an object if he's constitutionally incapable of having it. Hence a person's practical good must be something that he's capable of being moved to pursue for his own sake; and his well-being must consist in things that he's capable of caring about at all for his own sake. The version of internalism that's suggested by Railton's story thus emerges as an instance of the principle that 'ought' implies 'can'.

I'll admit that deriving internalism from the principle that 'ought' implies 'can' may seem like damning it with a faint premise. The rule of inferring from 'ought' to 'can' is highly controversial, and the grounds on which this rule was accepted by moralists of an earlier generation have been widely criticized.¹⁶ But the case for internalism does not require the principle that 'ought' implies

¹⁶ See, e.g., Stocker (1971), Sinnott-Armstrong (1984). For an excellent critical survey of the views of Sidgwick, Ross, Prichard, and others, see Frankena (1950).

'can' always and without exception. All it requires is a weaker principle, which I would formulate by saying that 'ought' can imply 'can'.

This weaker principle reflects the fact that inability sometimes is the reason why a person isn't obligated to do something. Arguments over whether inability always entails a lack of obligation have not undermined the intuition that it sometimes does. No one has yet denied, for example, that the reason why a person is never obligated to perform acts of telepathy or telekinesis or levitation is precisely that he can't.¹⁷

I am not going to present a detailed account of how and when inability entails a lack of obligation. Any such account would require fine discriminations among kinds of obligation, on the one hand, and kinds of ability, on the other, followed by an assessment of the connections obtaining among their various permutations.¹⁸ What can be said in advance, I think, is that obligatoriness must attach to things that are options, in some sense of the word, and that something's being impossible threatens its status as an option. The property of being obligatory is that of bearing a particular status in the context of practical reasoning; and if something isn't an option, it may be excluded from that context and hence ineligible to have any status within it.

9. *'Prima Facie Ought' and 'Prima Facie Can'*

These general (and admittedly vague) remarks imply nothing about the substantive issues that philosophers have sought to resolve with the principle that 'ought' implies 'can' – issues such as whether a person's obligations can ultimately conflict, or whether ignorance can exempt a person from duty. Yet they do enable us to make some headway on the question at hand, which is whether internalism about a person's good may ultimately rest on some version of the principle that 'ought' implies 'can'.

We can note, to begin with, that insofar as "good" means "such as one ought to want" or "... pursue" or "... care about", the occurrence of 'ought' in the *definiens* expresses a prima facie obligation. Some things that are good for a person may not, in the final analysis, be such as he ought to pursue or even care about for his own sake, since the final analysis may encompass competing and more pressing claimants for his attention. That something is good for a person entails only that he ought to pursue or care about it in the first instance, other things being equal.

¹⁷ But suppose that I promised you that I would read someone's mind. Am I not then obligated to do so? I think not, although I am of course guilty of having made a lying promise.

¹⁸ Again, I recommend Frankena's (1950) as a prolegomenon to such a project.

Some philosophers have argued that *prima facie* obligations fall outside the intended range of the principle that 'ought' implies 'can'.¹⁹ But William Frankena suggests that *prima facie* obligations do require latent ability, if nothing more.²⁰ I think that a case can be made for a version of Frankena's position.

That case begins with the observation that the concept of *prima facie* status can be applied to options as well as to obligations. Some things that are impossible in the final analysis are nevertheless *prima facie* options, in the sense that their impossibility isn't settled in advance of practical reasoning about whether and how to undertake them. Impossibility will eventually exclude these things from practical evaluation, but their exclusion will consist in their being thrown out of court, so to speak, rather than in their never being admitted in the first place. Hence their impossibility isn't such as to exclude them from consideration in some initial phase of deliberation, or from having some status within that context. Being *prima facie* options, then, they are eligible to be *prima facie* obligatory. Other things, however, are impossible in ways that exclude them from even the most preliminary consideration, and these things are not even *prima facie* options. They include actions of which the agent is constitutionally incapable, things that are not within the behavioral repertoire of his species or kind. Such things cannot be obligatory even *prima facie*.

I am thus prepared to hypothesize that whatever isn't at least a *prima facie* option cannot be even *prima facie* obligatory.²¹ And this hypothesis entails that something cannot be good, in either sense, for a creature who is constitutionally incapable of caring about it self-interestedly, in the corresponding way.²² For if a creature is incapable of caring about something, then caring about it is not a *prima facie* option for him, and so he cannot be under even a *prima facie* obligation to care about it. And unless the creature ought *prima facie* to care about the thing self-interestedly, it isn't intrinsically good for him. Thus can a plausible version of internalism about the good be derived from a plausible version of the principle that 'ought' implies 'can'.

¹⁹ See Stocker (1971), pp. 308–10.

²⁰ Frankena (1950), p. 175.

²¹ Note, once again, that this assertion has no bearing on the question whether obligations can ultimately conflict, since it is about *prima facie* rather than ultimate obligations.

²² Similarly, Frankena concludes that 'ought' implies "is latently able" in the case of our obligations to have particular feelings (1950, p. 161). Note that I shall henceforth dispense with the distinction between a person's practical good and his welfare.

10. *What This Derivation Implies*

This derivation of internalism about the good reorients the doctrine on the metaethical map. It suggests, for one thing, that the plausibility of internalism about the good isn't due to any general connection between normativity and motivation or affect. Rather, the plausibility of internalism about the good is now traced to a peculiarity in the normativity of value – namely, that value is normative specifically for motivation and affect.²³ Being good is being valuable, and being valuable is being such as one ought *prima facie* to value or care about. That goodness is normative fundamentally for affect, in this sense, may be the only reason why it is contingent on a creature's affective capabilities.

Thus, the plausibility of internalism about the good may provide no support to internalism about the right, since the two doctrines may not be relevantly analogous. Rightness is not normative for affect in the first instance: there is no term of affective obligation that's synonymous with "right", as "valuable" is synonymous with "good". There is consequently no reason to expect the normativity of rightness to bear any connection to affect – not, at least, if the connection in the case of goodness is mediated by the principle that 'ought' can imply 'can'.

Another implication of deriving internalism from this principle is that our current, weak version of internalism about the good may be the strongest version that's true. In particular, we should not expect to find any truth in those versions which say that something's being good for a person depends on his having a positive disposition to care about it under conditions specifiable without reference to the good, such as his being rational and fully informed about nonevaluative matters.²⁴ What's prerequisite to a thing's being good for a person is simply that which is prerequisite to its being such as he *prima facie* ought to care about for his own sake; what's prerequisite to its being such as he *prima facie* ought to care about is that his caring about it be a *prima facie* option; and his caring about it is a *prima facie* option so long as caring about it is possible for a creature like him – meaning, I take it, so long as there are some conditions or other under which a creature like him would care about it. Even if he wouldn't care about the thing under one set of conditions, such as his being fully informed and rational, there may yet be other conditions under which he would, and so the thing may satisfy the only prerequisite to possessing value.

²³ The idea that value is normative for affect in the first instance, and for action only secondarily, is central to the work of Elizabeth Anderson. See Anderson (1993).

²⁴ Railton favors this version of internalism about the good. See note 3, above.

This implication accords with arguments that have recently been raised against the conventional wisdom about the relation between value and affect. Michael Slote has pointed out that a person is rationally permitted, and may even be rationally obligated, not to desire some things that are intrinsically good for him and that he even recognizes as such.²⁵ The reason, Slote argues, is that moderation in desires – wanting to have enough of what’s good rather than wanting to have it all – is rightly considered to be a virtue and must therefore be compatible with, or indeed a component of, rationality in desire. Thus, what a person would desire if he were fully informed and rational may not include everything that’s intrinsically good for him.

11. Conclusion

Of course, I have not proved that the plausibility of internalism about the good is in fact due the connection between ‘ought’ and ‘can’. I’ve only offered this derivation as an hypothesis, which proves nothing; but it does shift various burdens of proof. Those who wish to reject internalism altogether should now be prepared to explain why the connection between ‘ought’ and ‘can’ shouldn’t lead us to believe that in order for something to be good for a person, it must be something that he’s constitutionally capable of caring about self-interestedly – either by desiring it, in the case of his practical good, or at least by being glad about it, in the case of his well-being. And those who wish to adopt a broader or stronger version of internalism should not appeal to the intuitive plausibility of internalism, unless they are prepared to explain why that plausibility should not be attributed to the connection between ‘ought’ and ‘can’, which doesn’t support internalism about the right or strong internalism about the good.

References

- ANDERSON (1993). Elizabeth Anderson: *Value in Ethics and Economics*, Cambridge, Mass., 1993.
- BLACK (1950). Max Black (ed.): *Philosophical Analysis*, Ithaca 1950.
- BRANDT (1982). Richard B. Brandt: “Two Concepts of Utility”, in Miller/Williams (1982).
- BRANDT (1989). Richard B. Brandt: “Fairness to Happiness”, *Social Theory and Practice* 33 (1989).

²⁵ Slote (1989).

- DARWALL (1983). Stephen L. Darwall: *Impartial Reason*, Ithaca 1983.
- ELSTER (1983). Jon Elster: *Sour Grapes*, Cambridge 1983.
- FALK (1986). W. D. Falk: *Ought, Reasons, and Morality*, Ithaca 1986.
- FRANKENA (1950). William K. Frankena: "Obligation and Ability", in Black (1950).
- FRANKENA (1958). William K. Frankena, : "Obligation and Motivation in Recent Moral Philosophy", in Melden (1958).
- GRIFFIN (1986). James Griffin: *Well-Being*, Oxford 1986.
- HARE (1989). Richard M. Hare: "Brandt on Fairness to Happiness", *Social Theory and Practice* 59 (1989).
- MELDEN (1958). A. I. Melden (ed.): *Essays in Moral Philosophy*, Seattle 1958.
- MILLER/WILLIAMS (1982). Harlan B. Miller and William H. Williams (eds.): *The Limits of Utilitarianism*, Minneapolis 1982.
- PARFIT (1984). Derek Parfit: *Reasons and Persons*, Oxford 1984.
- RAILTON (1986). Peter Railton: "Facts and Values", *Philosophical Topics* 14 (1986).
- ROSATI (1995). Connie Rosati: "Naturalism, Normativity and the Open-Question Argument", *Noûs* 29 (1995).
- SIDGWICK (1907). Henry Sidgwick: *The Methods of Ethics*, seventh ed., Indianapolis 1981; first ed. first publ. in 1874, seventh ed. first publ. in 1907.
- SINNOTT-ARMSTRONG (1984). Walter Sinnott-Armstrong: "'Ought' Conversationally Implies 'Can'", *Philosophical Review* 93 (1984).
- SLOTE (1989). Michael Slote: *Beyond Optimizing*, Cambridge, Mass., 1989.
- STOCKER (1971). Michael Stocker: "'Ought' and 'Can'", *Australasian Journal of Philosophy* 49 (1971).
- VELLEMAN (1991). J. David Velleman: "Well-Being and Time", *Pacific Philosophical Quarterly* 71 (1991).