bott, 'Interpretations of Mill's *Utilitarianism*', *Philosophical Quarterly* 6 (1956), pp. 115-120, at p. 117.

[12] Mabbott, 'Interpretations of Mill's *Utilitarianism*, at p. 116.

[13] See R. M. Hare, *Moral Thinking: Its Level, Method, and Point*, Oxford: Clarendon Press, 1981, ch. 2 (esp. p. 43), also pp. 50, 155-156; and Henry West, An *Introduction to Mill's Utilitarian Ethics*, Cambridge: Cambridge University Press, 2004, p. 87.

FROM: SOUTHWEST PHILOSOPHY REVIEW, VOL. 23, NO. 1 (JAN 2007)

Internalism and the Self: Lessons from Korsgaard's Kantian Critique of Williams

Daniel Callcut

1: Two Forms of Skepticism

The main goal of Christine Korsgaard's "Skepticism about Practical Reason,"[1] is to argue against the Humean idea that skepticism about practical reason can arise from psychological considerations alone. Korsgaard distinguishes between two distinct forms of skepticism about practical reason: content skepticism and motivational skepticism. Korsgaard's aim, in her terms, is "to establish the fact that motivational skepticism has no independent force."[2]

Content skepticism, as the name suggests, concerns the scope or range of reason. The content skeptic about practical reason comes in two familiar forms: the first doubts that reason can offer more than instrumental guidance in the realm of action and the second, more radically still, doubts whether reason has anything at all to say about how we live our lives. Motivational skepticism, as the name suggests, concerns the power of reason to move us to action. Motivational skepticism, in parallel with content skepticism, comes in two forms: the first doubts whether reason is capable *by itself* of moving us to action and the second, more radically still, doubts whether reason, even with assistance, ever moves us to do anything. Both varieties of skepticism, content and motivational, call to mind Hume, where, Korsgaard notes, skepticism about practical reason receives its "classical formulation."[3]

Hume's skepticism about practical reason, Korsgaard suggests, is question-begging. Hume does not look at human action and then consider the extent to which the concept of reason is appropriately invoked in describing or criticizing it; rather, Hume approaches the issue with an antecedently fixed list of what counts as an operation of reason and, given this list, it is a foregone conclusion that reason alone will turn out to be practically impotent. Hume's content skepticism, his stringent list of the operations of reason, underpins his motivational skepticism, his view that reason can only influence action via the passions. Korsgaard argues that Bernard Williams' motivational skepticism in "Internal and External Reasons,"[4] while more subtle than Hume's, is ultimately just as dependent upon skepticism about the content of reason.[5]

2: Williams' Internalism

I will not rehearse all the details of Williams' well-known conception of reasons for action here.[6] At the core of Williams' internalist view is the following: for an agent to have a reason to act there must be something about his or her motivational state that gives him or her a reason for action. The effect of internalism, for Williams, is to place an agent's motivations at the centre of the very idea of a reason for action: to work out what reasons people have, the content of their reasons, we must look to their motivations, since, for any agent, these are the source of his or her reasons for action. Hence, once internalism is adopted, motivational considerations do seem to constrain content. If a man is ultimately devoid of any motivation to save a drowning child, for example, then he has no reason to save her, and that's that.[7]

Williams' view is clearly not of the strong form of motivational skepticism according to which reason never moves us to do anything, for whenever we are moved to action via one of the rational processes that he has specified (most obviously, when we follow the means to our end) then we are moved, in part, by reason.[8] However, for Williams, while the rational processes may single out ends for us, they can do so only in light of other ends. Thus, in order for the rational processes to get started, it seems that there must be some ends that are not generated via the rational processes and that are in themselves neither rational nor irrational.[9] More pointedly: if reason by itself *cannot* select any ends for action, then none of the motivations that anyone has *can be* the product of reason alone. And to arrive at this view simply is to arrive at the weaker form of motivational skepticism introduced above. Williams' view that reason alone can never move one to action thus depends, apparently, on his view about what practical reason is, which is to say, in a nutshell, an instrument for ordering and satisfying one's given motivations. In other words, his skepticism about the motivational power of reason depends, apparently, on an implicit view about the content of reason, which is exactly what Korsgaard claims. However, it can look, as Korsgaard also observes, the other way around, as if motivational considerations are themselves constraining the content of practical reason. This observation leads directly to Korsgaard's central point.

3: Korsgaard's Case

Korsgaard's central point against Williams in "Skepticism about Practical Reason" can be expressed as follows: Williams conflates the internalist thought that for an agent to have a reason to ϕ there must be something about her (something in her 'S', to use Williams' shorthand for a person's "subjective motivational set"[10]) that gives her a reason to ϕ with the instrumentalist thought that an agent has a reason to ϕ only if she has a preexisting *goal* in her S that can be rationally connected to ϕ-ing. It is the former thought that best expresses the spirit of what Williams has in mind by internalism. But Williams typically expresses, or seems implicitly to construe, the former thought in the latter instrumentalist manner. However, to do so is to beg a very important question. For Williams, in doing so, is simply assuming that there cannot be modes of reasoning that by themselves, not in the service of some preexisting goal, might motivate one towards or against certain kinds of action. If there are such modes of reasoning, then one would expect to find in someone's S, at least the rational person's S, principles of practical reason that rule in or rule out certain kinds of action. The categorical imperative is the most obvious example of such a principle. But one might also think that there are principles that are more directly substantive. For instance, one might expect to find a principle to the following effect: save the lives of people in grave danger close at hand when one can do so at a small risk to self. Williams, with his ontologically generous conception of what kinds of element might be found in someone's S, cannot, and without argument should not, foreclose the possible existence of such principles. But he does.[11]

4: Internalism and the Self

Korsgaard's paper helps us to see the vitally important connection between Williams' conception of reasons for action and his conception of the self to which reason-for-action ascriptions apply. But we can see Williams' case as one instance of an important general truth: that the issue of how to conceive of the self must be central for any internalist since all internalists are united in seeing the self as the locus of reasons.[12] If reason claims have to connect to something in the subjective motivational set of the person to whom they apply, then the key question must be: how should one conceive of this set? On the classic instrumentalist view, desires alone set the basic goals or ends of action and reason is and can only be the 'scout' or 'slave' that figures out the best way to achieve such desire-given ends. The extent to which people will have reasons for action in common, then, will depend precisely on the extent to which there is uniformity of desire among people. If one has a full-blooded account of human nature such as Hume's, then one can confidently judge the desires, and thus the reasons for action, of those in distant times and places. Williams, however, rejects any such strong account of human nature and emphasizes the variability of human cultures and traditions. The result is an account of the self, and its

reasons, according to which the content of both is a matter of considerable contingency and unpredictability. We cannot be sure what motivations, and thus what reasons, other people have. Indeed, the further away we move from our local community and its (perhaps) familiar goals and assumptions the less sure we might become.[13] Thus we see the force, in Williams' hands, of the internalist idea that what you have reason to do depends on how things are with you. Reason claims have to find something to hook on to in their target, and how can we presume, ahead of time, to know the condition of our target?

Korsgaard wants to argue for a different conception of the target. Her point is that Williams' liberal policy[14] regarding the possible contents of someone's S makes available, as it were, more resources in the self for reasons for action to find a home.[15] For the Kantian, rationality's place in the human self is not merely as a servant but as a source of our reasons for action. Indeed, for the Kantian, there are principles of practical reason that by themselves are capable of directing and motivating action.

5: Korsgaard's Kantian Internalism

Korsgaard sets out what she calls the "internalism requirement" as follows:

> "Practical-reason claims, if they are really to present us with reasons for action, must be capable of motivating rational persons."[16]

This is not an easy claim to interpret. Part of the problem stems from a Kantian ambiguity surrounding the term "rational." This can be taken in the common, strict normative sense, according to which we are not always rational. Or it can be taken to indicate a certain kind of being – human as opposed, say, to amoeba. We could be said always to be rational in this latter sense though not in the former (we are not always, as one might say, fully rational). At times, Korsgaard clearly has in mind the former, strict sense, as when she writes that internalism requires only that "rational considerations succeed in motivating us *insofar* as we are rational."[17] Partly this is because Korsgaard is concerned to combat the unfortunate idea that internalism requires that reasons always motivate those to whom they apply.[18] But it is also because Korsgaard does sometimes seem to see internalism as primarily a thesis about the relationship between fully rational agents and reasons for action, namely, as the quotation suggests, the thesis that *fully rational* agents are always moved by reasons. However, it seems to me that Korsgaard is best understood to be talking about rational beings only in the weaker sense in her formulation of the internalism requirement. For note that the requirement states that reasons must be *capable* of motivating

rational persons and, if Korsgaard meant the stronger sense of "rational" here (that is, fully rational), then, given her view that reasons *will* succeed in motivating fully rational beings, why does not she not say that if a consideration is to count as a reason that it *must* motivate rational beings? So, if the reference to "capable" isn't to be otiose, then she must be concerned to relate reason claims to imperfectly rational creatures like us.

There are deeper reasons for thinking that Korsgaard has, at least on some occasions, only the weak sense of "rational" in mind. Williams' concern is to specify what *gives* us our reasons, not how (if we are fully rational) we will act upon them, and Korsgaard does apparently wish to engage with this aspiration to say *what it is about us* that gives us our reasons. But, as Korsgaard is at pains to point out, we still have reason to do things even when, being less than fully rational, we fail to do them. So it cannot be that we only *have* reasons when we *are* fully rational. This means that if Korsgaard is claiming that our status as rational beings is the source of our reasons, as indeed she seems to be, then she must have the weak sense of "rational" in mind. This is not to say that the notion of full rationality is not involved in Korsgaard's conception of what we have reason to do, for I take it that she sees the two senses of "rational" as connected: that what makes us count as rational beings in the weaker sense is our capacity to be rational in the stronger sense. Hence it is Korsgaard's view, as I see it, that our capacity to be fully rational is, in some sense, the source of our reasons. The explanation for this stems from the fact that Korsgaard shares Williams' concern that reasons must be capable of motivating us. For Korsgaard could simply define what we have reason to do in terms of what a fully rational being would do and stop there. But if reasons are to be understood in terms of what a fully rational being would do, and *they are not to be external to us*, then we must all have the capacity to be fully rational. The capacity to be rational is thus what connects us to the standards of full rationality and gives us internal reason to act in accordance with them. Inside every imperfectly rational being there must be, as it were, a perfectly rational one waiting to get out.[19]

6: The Cost of Korsgaard's Gain

The question still arises, however, as to whether Williams' concern to have psychological considerations *constrain* what we have reason to do has been abandoned. For Korsgaard's position suggests a picture in which we should try to conform our psychology to normative principles rather than the other way around. On this picture, reasons are seen as standing over our psychology: "A person with good character will be, on such a view, one who responds to the available reasons in an appropriate way, one whose

motivational structure is organized for rational receptivity."[20] Thus, while it is Williams' view that the basic ingredient of reasons for action is the contingent orientations of our psychology, which are then organized by rational principles, Korsgaard's Kantian view is that pure rationality alone is capable of giving us direction. Korsgaard does suggest in places that she shares Williams' view that motivational considerations constrain normative principles. She writes, for instance, that "the internalism requirement...does not refute ethical theories [but that it] *makes a psychological demand on them.*"[21] This suggests that motivational concerns have a real bite. However, Korsgaard surely renders any such bite toothless by proposing that the internalist connection between reason and motivation can always be preserved by transcendental argument: if one *assumes* that reasons must be capable of motivating us, then whenever one thinks that we *have* reason to act in a certain way one must conclude that we are *capable* of acting in the manner that these reasons demand.[22]

What this reveals is a conception of internalism in which normative principles are used to place demands upon our psychological capacities rather than the other way around.[23] Internalism, viewed this way, turns out to be no constraint on normative theorizing whatsoever. Whatever normative principles are taken to apply to everyone are automatically incorporated into everyone's S. This enables the Kantian, in stark contrast to Williams, to secure a priori knowledge of universal reasons for action. But the price of doing so must surely be, contrary to what Korsgaard suggests, the abandonment of Williams' aim of placing psychological constraints upon normative theory: for if one is to capture the intuition that motivational considerations place real constraints on what someone has reason to do then one cannot help oneself, in determining what a person has reason to do, to whatever psychological resources one needs.

It is hard to deny Williams' concern to grant in sober fashion the possibility of moral claims finding no psychological resonance whatsoever in those to whom they apply *if* one gives up Kant's transcendental psychology and embraces the broadly 'empirical' conception of the mind that is now dominant. Here I am in sympathy with Gideon Rosen's claim, made in another context, that Kant's conception of the mind is simply not a live option for us anymore.[24] As such, I think we have to agree with Williams that the possibility of people being left cold by moral claims cannot be ruled out on a priori grounds. Indeed, Williams would rightly urge, we know that this is no mere abstract possibility but all too often a sad and nasty reality. This, of course, does *not* mean that Williams is right to foreclose the possibility that people *can* act in accordance with principles that by themselves rule in

or out certain kinds of action. The point is that he is right to argue against the view that such principles are necessarily present in all people.

However, if we keep Williams' commitment to an empirical approach to human behaviour and motivation and we combine it with internalism, then it now becomes (in part) an empirical matter as to whether people have reason to save or not to save drowning children, or to kill or not to kill innocent bystanders, and so on.[25] If we wish to preserve a robust sense of moral reasons, this is a serious problem. My view is that the correct response to the problem is not to abandon (at least in Rosen's generous sense) an empirical conception of human motivation. The difficulties with both Williams' and Korsgaard's positions should lead us to consider rejecting the premise they share: the internalism that requires one, in too strong a way, to tie the theory of normative reasons for action to a theory of the self. But that is a story for another day.[26]

Notes

[1] Christine Korsgaard, "Skepticism about Practical Reason" in *The Journal of Philosophy*, Vol. LXXXIII, No.1, Jan 1986, 5-25. All references to Korsgaard are to this paper.

[2] Korsgaard, 6.

[3] Korsgaard, 6.

[4] Bernard Williams, "Internal and External Reasons" in *Moral Luck* (Cambridge: Cambridge University Press, 1981), 101-113. All references to Williams are to this paper.

[5] Korsgaard invites us to notice a further point, namely that Hume's list of the operations of reason also determines the potential candidates for the job of slave to the passions; and the only remotely qualified candidate on the list is that of means-end causal reasoning. Korsgaard points out: "One could disagree with Hume about his list of the types of rational judgement, operation, or possible deliberation, and yet still agree with the basic point about the source of motivation: that all rational motivation must ultimately spring from some nonrational source, such as passion" (8). This is precisely what Korsgaard takes Bernard Williams to do in "Internal and External Reasons." John Skorupski argues in an important forthcoming paper, "Internal Reasons and the Scope of Blame," that the relationship of Williams' position to that of Hume is complex and that Williams' view includes many non-Humean elements.

[6] The basic idea is as follows. There are two different sorts of interpretation of reason ascriptions: internal and external interpretations. According to an internal reading of reason ascriptions, there is something about the agent, about his or her aims, that *gives* him or her reason to ϕ. If the agent has no such aims, then he has no such reason to ϕ, and the statement is false. By contrast, according to an external interpretation of such reason ascriptions, the fact that none of an agent's

standing goals would be furthered by ϕ-ing does not necessarily entail that such an ascription is false. An externalist reading of such statements allows, then, for the possibility that an agent has reason to ϕ regardless of his or her aims, dispositions, and so on. We can call, as Williams does, the reasons that internal interpretations ascribe internal reasons; the reasons that external interpretations ascribe, external reasons. Williams' central claim in "Internal and External Reasons" is that only internal reasons exist.

[7] For more discussion of this point, see Chapter One of my doctoral thesis, *Bernard Williams and the End of Morality* (Baltimore: Johns Hopkins University, 2003).

[8] Williams provides examples of rational processes on 104. They key point for what follows is that they are all instrumental in nature.

[9] For interesting discussion of this point in relation to foundationalism, see Joseph Heath's "Foundationalism and Practical Reason" in *Mind* Vol. 106, no.423, 1997, 451-473.

[10] Williams, 102.

[11] For extensive (and illuminating) discussion of this point, see Joshua Gert, *Brute Rationality: Normativity and Human Action* (Cambridge: Cambridge University Press, 2004), chapter 3. Korsgaard does not think that Williams is, strictly speaking, begging the question. She writes: "Although I think [Williams] is quite right in saying that the burden of proof about what is to count as a purely rational process – about content – belongs to Hume opponents, I am arguing that there is no reason to suppose that if this burden is successfully picked up the [resulting conception of] reasons will be external" 22, n. 16. Thus her concern is rather to indicate that Williams has run together questions about the content of reasons and the question of whether reasons must be internal or not. Actually, Korsgaard's way of putting the point is unfortunate, for it seems to suggest that an account of reasons for action must fall out of an account of rationality whereas an important question is whether one should think of reasons as derived from rational procedures at all.

[12] Jonathan Lear anticipates this point in his marvelous paper, "Moral Objectivity" in *Objectivity and Cultural Divergence* edited by S. C. Brown (Cambridge: Cambridge University Press, 1984), 135-170.

[13] I do not mean to imply here that people always are more knowledgeable about those around whom they live: indeed, superficial familiarity can of course generate an undeserved confidence.

[14] Williams rejects the view that the only elements to be found in an agent's subjective motivational set are desires. As he says, the term 'desire' can be used, formally, to refer to all the elements in S. However, he continues, "this terminology may make one forget that S can contain such things as dispositions of evaluation, patterns of emotional reaction, personal loyalties, and various projects, as they may abstractly be called, embodying commitments of the agent" (105).

[15] Now it is worth stressing that Williams' instrumentalism is surely not, nor perhaps does Korsgaard mean to suggest, merely the residue of a commitment to a 'belief-desire' theory of action since, after all, one can adhere to the 'belief-desire'

theory without being an instrumentalist about rationality. But Williams' liberal policy is significant in that it opens the door for a Kantian form of internalism that simply does not show up on Williams' philosophical radar. This is the Kantian position that Korsgaard aims to defend. It would seem that this different form of internalism, given the above thought that the notion of the self must be central for any internalist, should give us a different conception of the self. And, indeed, as we shall see, it does.

[16] Korsgaard, 11.

[17] Korsgaard, 15, emphasis added.

[18] There is no need, she rightly insists, for the connection between reasons and motivations to be that strong. Indeed, the absurd idea that it must be is behind much mischief, in particular the idea that moral reasons must have a kind of magic attractiveness. Thus if one can say "So what?" to a putative reason or putative obligation, the idea seems to be, then it cannot really be one. But the fact that the norms of morality can fail to be observed is as good a basis for moral skepticism, at least at first blush, as the fact that the norms of logic can fail to be observed is a basis for skepticism about logic. This shows that internalism and the "So what?" argument, as it could be called, needs to be handled carefully. Williams, it is worth stressing, can agree with this point: if you are rational, then you will act on your reasons (at any rate: those reasons that rationality *requires* you to act upon) but there is no guarantee that you *will* act upon them. There is no guarantee, in short, that you will be fully rational.

[19] This, of course, is a characteristically Kantian idea, given dramatic form in Kant himself through the distinction between our phenomenal and noumenal selves.

[20] Korsgaard, 8.

[21] Korsgaard, 23, emphasis added.

[22] I should note that I am describing Korsgaard's view as I read it in "Skepticism about Practical Reason." Korsgaard's subsequent work shows far greater concern to avoid such a 'stipulative solution', though it arguably ends up jeopardizing the moral universalism as a result.

[23] Korsgaard's view, in this regard, is the same as Nagel's, whose view she describes and endorses as follows: "the internalism requirement leads not to a limitation on practical reason, but to a rather surprising increase in the power of moral philosophy: it can teach us about human motivational capacities: it can teach us psychology" (23). The work of Nagel's alluded to is *The Possibility of Altruism* (Princeton, N.J.: Princeton University Press, 1970). Nagel's aim is the Kantian one, shared by Korsgaard, of grounding the universal applicability of normative principles (more specifically, ethical principles) in aspects of human psychology that are (so it is claimed) universal. He gives eloquent expression to this aspiration as follows: "[M]y position resembles Kant's in that it assigns a central role in the operation of ethical motives to a certain feature of the agent's metaphysical conception of himself. On Kant's view the conception is that of freedom, whereas on my view it is the conception of oneself as merely a person among others equally

real. However, different as they are, both are thought to be conceptions which we cannot escape, and are thought to provide that basis for ethical motivation which in other internalist theories is provided by various motives and desires. Because of the alleged inescapability of these conceptions, a view of the Kantian type entails that we are not fully free to be amoral, or insusceptible to moral claims. That is what makes us men" (14).

[24] Gideon Rosen, "Objectivity and Modern Idealism: What is the Question?" in *Philosophy in Mind* (Dordrecht, The Netherlands: Kluwer Academic Publishers, 1994), ed. M. Michael and J. O'Leary-Hawthorne, 277.

[25] Williams, not implausibly, stresses how rare it will be that someone has no reason, on his account, to pursue his or her well-being, and perhaps the same case could be made for moral motivation. Still, it is also worth stressing that the question of whether someone has such reason now depends on psychological contingencies. Whether you have reason to be prudent on this occasion or that, or moral on this occasion or that, depends on the empirical state of your psychology. This point is developed at length in Skorupski's forthcoming paper though he does not draw the anti-internalist moral from it that I do.

[26] I tell part of the story in Chapter One of my doctoral thesis, *Bernard Williams and the End of Morality* (Baltimore: Johns Hopkins University, 2003). I would like to thank Joshua Gert, Vaughn Huckfeldt, and Jerry Schneewind for the many long and illuminating conversations we have had on the issues discussed in this paper.