

Joyce considers the example of daily workout:

Suppose I am determined to exercise regularly, after a lifetime of lethargy, but find myself succumbing to temptation. An effective strategy will be for me to lay down a strong and authoritative rule: *I must do fifty sit-ups every day, no less*. I am attempting to form a habit, and habits are formed – and, for the doggedly weak of will, maintained – by strictness and over compensation. Perhaps in truth it doesn't much matter that I do fifty sit-ups every day, so long as I do more-or-less fifty on most days. But by allowing myself the occasional lapse, by giving myself permission sometimes to stray from the routine, I pave the way for akratic sabotage of my calculations – I threaten even my doing more-or-less fifty sit-ups on most days. I do better if I encourage myself to think in terms of fifty daily sit-ups as a non-negotiable value, as something I must do if I am ever to get fit.¹⁴

The proposition maintaining the absolute necessity of fifty sit-ups is transparently false yet supposedly best affirmed in the modus of make-belief, according to Joyce. What he describes is a real psychological phenomenon most of us are struggling with. What I hesitate to grant, however, is that any transparently false make-belief is better positioned than a clear-headed alternative true belief to ameliorate the present practical quagmire.

In fact, it seems to me that what Joyce describes is precisely what we should believe, and all that we should believe (or make-believe), since it seems to deliver the same practical benefits. It is true that Joyce, determined to exercise, needs to lay down a strict law permitting no exceptions. But the attitude he has no need for is that of falsely make-believing that only fifty-sit-ups could serve his athletic purposes. Instead, when facing temptation, what he needs to remind himself of is that breaking the law considerably weakens his habits, compromising his future resolve to do any exercise, and that as a psychological matter, he must fully abide by his strict workout plan if he is to stand any chances of sticking to it. Nothing in that realization is mistaken, supposedly requiring a fiction to maintain it, but rather represents a clear-headed awareness of the psychological implications of what we do now for what we may do later as well as our general dispositions and habits. Placing our current choices in a larger psychological context, appreciating their long-term effects among other things, endows them with considerable gravitas, quite possibly outweighing the kind of hallucinatory gravitas knowingly absent in reality and present only in our fictional fantasies.

14 R. Joyce, *Moral Fictionalism*, p. 303.

Review Article

∴

Christine Korsgaard's *Self-Constitution*

Terence Cuneo and Randall Harp

The Department of Philosophy, University of Vermont

tcuneo@uvm.edu; randall.harp@uvm.edu

Abstract

Christine Korsgaard's 1996 book, *The Sources of Normativity*, attracted a great deal of attention. And rightly so. It is a highly engaging attempt to answer what she calls the normative question, which is the question of what could justify morality's demands. Korsgaard's latest book, *Self-Constitution*, develops and defends the broadly Kantian account of action and agency that hovers in the background of *Sources*, drawing out its implications for the normative question. In this review, we present the main lines of argument in *Self-Constitution*, raising objections to both Korsgaard's account of action and agency and her most recent attempt to address the normative question.

Keywords

action – agency – authority – categorical imperative – hypothetical imperative – Kant – normative question – realism – self-constitution

Christine Korsgaard, *Self-Constitution* (Oxford: Oxford University Press, 2008), 230 pages. ISBN: 978-0-955279-5 (hbk). Hardback/Paperback \$105/35.00.

Christine Korsgaard begins her book *The Sources of Normativity* (Cambridge, 1996) by noting that what morality requires is hard. For morality, Korsgaard

observes, makes extensive and radical demands of us, sometimes requiring even that we surrender our lives. What could justify these demands? This last question is what Korsgaard calls “the normative question” and her project in *Sources* is to answer it. Those familiar with Korsgaard’s discussion of the normative question know that it is encyclopedic, drawing upon figures as diverse as Plato, Aristotle, Hume, Wittgenstein, and, most of all, Kant. There is, however, a unifying element lying in the background of Korsgaard’s wide-ranging discussion: a particular view of both action and agency according to which an agent can act only if she “identifies” with a principle of choice, such as the categorical imperative (*Sources*, p. 246).

Korsgaard’s most recent book, *Self-Constitution*, is an attempt to develop and defend this Kantian account of action and agency. Many of us have anticipated this book with great interest. The reason why is that it is easy to feel deeply ambivalent about *Sources*. On the one hand, the book is brimming with insight, exhibiting philosophical imagination and creativity of the highest order. Who would’ve thought that we could blend insights from Plato, Kant, and Wittgenstein into a coherent normative view? On the other hand, when one unpacks the book’s arguments, they appear to not even come close to offering a satisfactory answer to the normative question. Indeed, the book appears to set itself up for failure, for central to Korsgaard’s project is the assumption that a satisfactory answer to the normative question must imply that failing to act as morality requires would be catastrophic, indeed, *incoherent*. This, however, seems like a standard that no normative theory could meet, including Korsgaard’s. The question that many of us have had is whether *Self-Constitution* would do better.

The answer is that *Self-Constitution* does, in various respects, do better. It addresses many of the concerns that one might have about the details of a broadly Kantian view of action and agency, such as the concern that the Kantian view cannot explain the fact that non-human animals act. But, in other respects, *Self-Constitution* leaves things more or less how we found them in *Sources*. There are no substantial alterations in the way that Korsgaard either frames or addresses the normative question. Accordingly, if one had worries about the basic approach taken in *Sources*, these worries will probably remain intact. Our project in this review essay is to summarize the book’s overarching argument, filling in at various points important details regarding Korsgaard’s Kantian account of action and agency. We then raise two sets of worries, the first of which concerns the Kantian account of action and agency developed in *Self-Constitution*, the second of which

concerns Korsgaard’s attempt to deploy this account to address the normative question.

I

Korsgaard’s overarching argument comes in three stages. Stated in skeletal form, it runs as follows. In the first stage, Korsgaard claims that human beings and non-human animals are agents, since they necessarily perform actions (Korsgaard declines to elaborate on the type of necessity involved, other than to say that it is not causal, logical, or rational; see pp. 1–2). To perform actions, says Korsgaard, is to be subject to normative standards. Which sorts of normative standards apply to agents? That depends on the type of agent in question. Non-human animals, which act from instinct, are subject to standards of both efficacy and autonomy. These standards require both that they adopt means that successfully produce their ends and that their actions constitute their Aristotelian form, contributing to what it is to be an animal of their kind. Human animals are also subject to standards of efficacy and autonomy. However, these standards are different from those that apply to non-human animals insofar as they are rules that govern not instinct but free choice (pp. 107–08).

In the second stage of the argument, Korsgaard argues that the standards of efficacy and autonomy, as they apply to human agents, are in fact the hypothetical and the categorical imperative. These two imperatives (which Korsgaard claims are simply different manifestations of one imperative) are *constitutive* of the actions of agents endowed with rational capacities: to fail to act in accordance with them implies a failure to be an agent of this sort. It follows from this, says Korsgaard, that insofar as human beings are agents, they must take the hypothetical and the categorical imperative to be authoritative with respect to their deliberation and action.

The third stage of Korsgaard’s argument is that the hypothetical and the categorical imperative (in its universal law formulation) entail not simply requirements concerning how to be effective practical agents, but also general *moral* obligations. Or, as Korsgaard puts it, they imply requirements of both *inward* and *outward* justice. Inward justice is a matter of how we treat ourselves—how we deliberate, what we take to be our ends and our means, etc. It is achieved when an agent achieves a high degree of coherence among her mental states. Outward justice is a matter of how we treat others. It is achieved when we treat others as ends in themselves. The hypothetical imperative and categorical imperative commit us to pursuing both inward and outward justice.

II

Let us now fill in some of the details of Korsgaard's overarching argument, beginning with her theory of action. Fundamental to Korsgaard's theory of action is a distinctive understanding of the relationship between *agents* and *actions*. To appreciate the view's distinctiveness, it is helpful to contrast it with an alternative way to understand action, which we shall call the *standard view*.

According to the standard view, there are *agents* and there are *behavioral events*. Proponents of the standard view offer a variety of accounts of what it is to be an agent. According to some of its advocates, agency consists in a thing's possessing certain characteristic mental states such as intentions. According to others, agency consists in a thing's being responsive to reasons. According to yet others, agency consists in possessing a will. While advocates of the standard view offer different accounts of agency, they generally agree about what it is for something to be a behavioral event. The standard view's position regarding behavioral events is (roughly) that every change in the world that is identified with or produced by an agent (or some part of an agent) is a behavioral event. This way of framing things generates the so-called problem of action, which is the task of determining which behavioral events count as actions, whose actions they are, and why. Take, for example, a case in which you stand on a street corner and wave your arm. Is your waving your arm an action and, if so, is it your action? Is the taxi cab stopping after you wave your arm an action and, if so, is it your action? Is your yawning as you enter the cab an action and, if so, is it your action?

The standard view, then, attempts to explain what counts as an action in terms of some relation that agents bear to behavioral events. One way to understand Korsgaard's claims about action in *Self-Constitution* is as a rejection of the standard view. For Korsgaard does not take the fundamental issue facing action theorists to be which behavioral events count as actions; rather, she takes it to be what relationship actions bear to agents.

Indeed, as Korsgaard sees things, the relationship that actions bear to agents is the reverse of that defended by advocates by the standard view. Rather than explain action in terms of agents, we explain agency in terms of actions:

The intimate connection between person and action does not rest in the fact that action is caused by the most essential part of the person, but rather in the fact that the most essential part of the person is *constituted* by her actions (p. 100, emphasis original).

According to this view, something is an action, and indeed a particular agent's action, just in case that action is such as to constitute that agent. But what,

according to Korsgaard, is the proper account of action? And what is it for something to be constituted by its actions?

An action, says Korsgaard, is "an act done for the sake of an end" (p. 11). As Korsgaard thinks of them, acts are the behavioral components of an action, such as waving one's arm or making a false promise. Ends are that for which one acts. They include states and events such as *flagging a cab* and *getting some ready cash* (p. 12). A consequence of this definition of actions is that actions (as opposed to acts) do not produce ends. Rather, ends (like acts) are components of actions. As such, says Korsgaard, it is not ends but actions that are the objects of choice. Note that, under this understanding of actions, acts themselves must have an intentional component, since making a promise which turns out to be false through no fault of the agent is a different act from making a promise which one has no intention of keeping. This account of acts separates Korsgaard's position from the broadly Davidsonian view that acts themselves are, properly speaking, non-intentional behaviors or events, though they might also be actions in virtue of bearing a certain relation to the agent, such as being intended by the agent (or being an ingredient in some intentional description that links the agent to the event).

Actions, then, not only causally connect agents to ends, but also constitute agents. What does it mean for an action to constitute an agent? Early in her discussion, Korsgaard tells us that the "task of self-constitution"—at least when it comes to humans—"involves finding some roles and fulfilling them with integrity and dedication. It also involves integrating those roles into a single identity, into a coherent life" (p. 25). By a self, then, Korsgaard has in mind not the sort of thing that forms the subject matter of discussions of personal identity in Anglo-American metaphysics, but rather a "practical identity," which she glosses as a "description under which you value yourself and find your life worth living and your actions to be worth undertaking" (p. 20). At no point in her discussion does Korsgaard mention Kierkegaard or other existentialists. But were one to think of what Korsgaard has in mind by self-constitution as what Kierkegaard calls self-creation, then one would not be too far off. Self-constitution is the project of identifying and orienting one's life around projects and values of one's choosing. When an agent acts in such a way that she endorses these projects and values, she identifies with them, thereby constituting her practical identity or self.

But how does self-constitution work? The first clue Korsgaard gives us is that all living things are engaged in an endless activity of self-constitution (2.4.1, 5.4.1). For non-human animals, self-constitution occurs by way of instinct. Non-human animals have self-maintaining Aristotelian forms, which lead them to aim at survival and persistence; it is their instincts that causally influence them to perform acts that aim at those ends. For human beings, by

contrast, self-constitution occurs by way of conforming to standards of rationality; it is not instincts but conformance to these principles that leads them to perform actions that aim at their ends. So, consider a case in which an agent faces an incentive, which Korsgaard says is “a motivationally loaded representation of an object” (pp. 104–05). If the agent is an animal, its instincts will cause it to react in such a way as to further its ends of survival and reproduction. In this scenario, instinct is the agent’s “contribution” to action (p. 92). If, by contrast, the agent is a human, rational principles determine how the agent is motivated to respond to the incentives. In Korsgaard’s view, these rational principles are the agent’s contribution to action.

As we noted earlier, these two standards of action—efficacy and autonomy—correspond in rational agents to the two Kantian imperatives, the hypothetical and the categorical imperative. The hypothetical imperative commands us to take the necessary means towards our ends; this corresponds to the norm of efficacy because it is only by taking the necessary means that we can successfully bring about our ends.¹ The categorical imperative, as mentioned, corresponds to the norm of autonomy. It governs the choice of actions by posing an admissibility test for acts being taken in pursuit of ends, evaluating both the act and the end together. Korsgaard also suggests that in satisfying the categorical imperative, we both make ourselves an agent and the cause of our desired end, where this causality is attributed to our own agency. The suggestion seems to be that when we satisfy the categorical imperative, we represent the principle that underlies our action as one which is rationally permissible for us *as agents* to act on, which means that we are willing to have that principle constitute us (p. 131). An implication of this position is that agency is, in Korsgaard’s view, episodic. When we violate the categorical imperative, as we often do, we fail to be agents.²

Both the hypothetical and the categorical imperative are thus norms that are constitutive of action, imposing normative obligations on agents. Still, says Korsgaard, we must explain how these norms actually bind us, giving us decisive reasons to act. Korsgaard argues that neither empiricism nor rationalism, which are the main competitor theories of action, can account for the

1 As it is typically understood, the hypothetical imperative is disjunctive, enjoining us either to take the necessary means toward our ends or to surrender those ends. Korsgaard does not emphasize the second disjunct of this injunction. Perhaps this is because thus understood the hypothetical imperative is not clearly a norm of efficacy.

2 Korsgaard says in several places that actions come in degrees (pp. 25, 163). However, it is difficult to see how this could be so, as something is an action, says Korsgaard, inasmuch as it satisfies the categorical imperative. But being such as to satisfy the categorical imperative is an on/off property; actions cannot satisfy the categorical imperative to some degree or other.

bindingness of these norms on action. The empiricist cannot account for norms at all, according to Korsgaard, because the empiricist has a purely descriptive account of action. The rationalist, by contrast, accepts the existence of normative facts, but cannot explain why those normative facts bind us.

Consider Korsgaard’s empiricist. According to the empiricist, when an agent is faced with a sufficiently strong desire, that desire suffices to cause her to act. There is, according to the empiricist, almost nothing more to say about what an action is. There is, for example, no need to appeal to norms of practical reason that guide the agent’s practical deliberation. But if this is the whole story regarding action, it is impossible, says Korsgaard, for our actions to fail to satisfy norms of practical reasoning and, hence, be practically irrational. The empiricist proposal fails to take into consideration the fact that an action requires that an agent’s beliefs and desires must be combined *in the right way* (4.1.3). If Korsgaard is right, this is not a problem that is easily fixed. The only way to solve it is to posit an agent whose task it is to combine beliefs and desires in accordance with the hypothetical and the categorical imperative so as to produce actions.

Now consider Korsgaard’s rationalist. Unlike the empiricist, the rationalist recognizes the hypothetical imperative as a rational requirement on action. The rationalist also recognizes that agents grasp this requirement. But the mere apprehension of this requirement, Korsgaard claims, does not explain the fact that agents are governed and guided by it. This requirement can govern and guide agents in their practical deliberation only if they *already have a prior commitment* to doing what rationality requires.

If these objections hit their mark, neither the empiricist nor the rationalist can account for the normative bindingness of the hypothetical and the categorical imperative. The empiricist’s account of action is defective because it is not normative at all, but purely descriptive. The rationalist’s position is inadequate because it fails to integrate rational requirements into practical deliberation and agency in the right way. If Korsgaard is right, however, the Kantian account of action is not only normative, but also integrates rational requirements into the production of action in the right way. For being an agent simply is to be one who is governed and guided by both the hypothetical and the categorical imperative. To the extent that a person fails to be guided and governed by these imperatives, she fails to be an agent.

III

We now move from exposition to evaluation, raising two concerns about Korsgaard’s account of action and agency. The first problem concerns the role

that both the hypothetical and the categorical imperative play in practical deliberation. As Korsgaard views things, agents use these principles to *test* rather than *weigh*: we test maxims to see whether they are permissible, rather than weigh them to see which is most in accordance with our reasons. Testing is all well and good, as it often helps us to determine whether a maxim satisfies some minimum threshold of rationality. Indeed, there are theories of action, such as satisficing theories, which imply that there are deliberative contexts such that any action that meets such a threshold is rational to perform. The problem is that it is very difficult to make do with *only* a testing principle; practical deliberation often requires weighing.

For instance, an agent might deliberate over more than one action, where each prospective action is permissible according to the categorical imperative. If it is possible for each action to be permissible, but for one action to be more rational than the others, then a simple testing model of deliberation will not suffice. To take a simple example, a person might deliberate about what to prepare for dinner in ordinary circumstances. One type of dinner might involve great risk of poisoning oneself and one's guests. Preparing this type of dinner is clearly not rationally permissible. Another type of dinner poses no risk of poisoning, is healthy, is sufficiently enjoyable, and is not too preparation-intensive. Preparing it is rationally permissible. However, a last type of dinner is not only safe, healthy, and easy to prepare but will also delight all who eat it. Preparing this dinner is rationally preferable, as it is the ideal meal to make given the circumstances. The problem with a simple testing model such as Korsgaard's is that it fails to provide the resources to distinguish between the latter two possibilities.

The second objection we would like to raise concerns Korsgaard's treatment of rival positions. We saw above that Korsgaard argues that realists (which she takes to be equivalent to rationalists) cannot account for the bindingness or rational authority of practical principles, such as the hypothetical imperative. It is worth quoting what she says at length:

The realist supposes that there are eternal normative verities of some sort—facts about which act-types or actions are right, or facts about what counts as a reason for what. How do we act on these verities? Apparently, by applying them in particular instances. We apply our knowledge that an action is right by choosing it... But notice that this sort of account could not possibly explain the normativity of the hypothetical imperative. We can see this by thinking about how it would have to work. The agent would have to recognize, as some sort of eternal normative verity, that it is good, or that it is required, or that there is a reason, to take the

means to his ends...How does he act on this recognition? How does he apply it to the case at hand? ... [W]e cannot explain how we are motivated to act on the hypothetical imperative, much less how we are bound by it, by appealing to the hypothetical imperative *itself*... The point is that the hypothetical imperative cannot be a normative truth that we *apply* in practice, because it is the principle *in accordance with* which we are acting when we apply truths in practice (pp. 64–65, emphasis original).

What should we make of this argument?

Well, suppose we begin by making some allowances. Let us grant, for the moment, Korsgaard's assumption that defending the authority of practical principles requires eliminating any skeptical doubts about whether such principles are authoritative (2.1.7). Let's also overlook the fact that no decent realist position would claim that for a norm to be authoritative for an agent, she must recognize it as such—indeed, must recognize it *as* "some sort of eternal normative verity." And, finally, let's set aside the fact that we cannot validly infer that "the hypothetical imperative cannot be a normative truth that we apply in practice," from the claim that "it is the principle *in accordance with* which we are acting when we apply truths in practice." If this inference were valid, after all, then by the same argument modus ponens could not be a truth that we apply when doing logic because it is a principle with which we operate when doing logic—a conclusion that does not follow.

The crucial question to raise about Korsgaard's treatment of realism is this: Why think that a realist would defend the authority of the hypothetical imperative in the way that Korsgaard claims? True, the realist holds that the hypothetical imperative is a mind-independent normative fact. But she, no less than the Kantian constructivist, can claim that conforming to the hypothetical imperative is constitutive of practical reasoning. There is nothing about the realist's ontology that suggests otherwise. Not only can realists claim this, they have in fact claimed something very much like this.

Consider, for example, what a realist such as Thomas Reid says about the principles of common sense. The principles of common sense, in Reid's view, are those propositions that all well-formed mature human persons must take for granted in their everyday lives on pain of suffering from serious normative defects, such as being incoherent, failing to treat like cases alike, and so forth. They are both mind-independent facts and at least some of them are constitutive of thinking and acting. It is true that philosophers such as Reid do not hold that "only formal principles," such as the hypothetical imperative "can be directly normative." And it is true that realists such as Reid deny that "our substantive principles must be derivable from formal ones if they are to be binding

on the will" (p. 46). While these differences are significant, they are compatible with the position that, according to realism, some formal normative principles are authoritative in the sense that Korsgaard wishes to defend and for roughly the same reasons that Korsgaard identifies. It follows from this that both realists and constructivists such as Korsgaard can (and do) accept claims about the constitutive status of normative principles. On this matter, there is no important difference between these views.

IV

We began this review by noting that *Self-Constitution* is an attempt to develop the account of action and agency that hovers in the background of Korsgaard's earlier book *The Sources of Normativity*. In *Sources*, Korsgaard acknowledged that there is an apparent gap in her argument. Having argued that since I am an agent who acts for reasons, I must value *my* humanity, Korsgaard also acknowledged that it wouldn't follow that I have to value *your* humanity. Korsgaard argued that this apparent gap could be closed. The way to close it, she claimed, is by observing that since reasons are public, they are also agent-neutral in the sense that they lack reference to any particular agent.³ This would imply, Korsgaard said, that if I valued my own humanity, I would have to value humanity wherever I find it.

In the penultimate chapter of *Self-Constitution*, Korsgaard considers a closely related problem. The alleged problem is that Kantianism is guilty of "empty formalism," for there is no way, in a sufficiently wide array of cases, to derive substantive moral content from a formal principle such as the categorical imperative. By employing a series of moves similar to that used in *Sources*, Korsgaard maintains that this challenge can be met. The argument she offers runs like this:

- (1) Necessarily, if an agent interacts with other agents, then these agents deliberate together (pp. 190, 210).
- (2) Necessarily, if agents deliberate together, then they must share reasons (p. 191).
- (3) All reasons are either private or agent-neutral (p. 195).
- (4) Necessarily, if agents share reasons, then the reasons they share must be agent-neutral, as "considerations that have normative force for *me* as well as you" (191–92, emphasis original). Each agent must treat each other's

³ Korsgaard uses the terms "private" and "agent-relative" and "public" and "agent-neutral" interchangeably (see p. 191). We will do so as well.

reasons as "having authentic normative standing ... simply because they are your reasons, respecting them "for their own sake" (pp. 201–02). The alternatives would be to treat your reasons as if there were nothing or simply tools for my own purposes (p. 196).

- (5) Necessarily, if the reasons agents share are agent-neutral, then these agents must treat the humanity of those with whom they deliberate as an end in itself (p. 192).

Given the plausible assumption that we do interact with one another, it follows that:

- 6) Necessarily, agents who deliberate together must treat the humanity of those with whom they deliberate as an end in itself.

Does this argument lay the empty formalism charge to rest? We believe not. Let us make two observations about it.

First, suppose we distinguish agent-neutral from third-person instrumental reasons. According to the received understanding, agent-neutral reasons are reasons that do not make essential reference to the agent for whom they are reasons. Agent-neutral reasons to perform some action, then, are reasons not simply for me to perform some action but for anyone relevantly situated as I am. By contrast, third-person instrumental reasons, let's stipulate, are considerations that favor an agent's performing some action but only because that action satisfies the purposes or desires of some other agent. For example, suppose that failing to put the top onto the toothpaste tube would make your spouse angry. And suppose that there is nothing else that favors your acting in this way. This would be a third-person instrumental reason for you to put the top back onto the toothpaste tube.

In her case for premise (4) of the argument stated above, Korsgaard seems to think that to establish that a reason is agent-neutral is thereby to establish that it is also not third-person instrumental. For, Korsgaard claims, if I were to treat your reasons "as private reasons, with normative force only for you," then I would see them simply "as tools or obstacles in the pursuit of my own ends" (p. 196). And one cannot do that and genuinely deliberate together. Rather, to deliberate together, each deliberator must treat each other's reasons as "having authentic normative standing ... simply because they are your reasons, respecting them "for their own sake" (pp. 201–02).

But consider a case in which we are both devotees of a third party, say, an Aztec goddess of some sort. Suppose, on some occasion, we believe that we have decisive reason jointly to perform a sacrifice to this goddess. The reason

in question is not agent relative or private. It is a reason for anyone situated as we are. Still, this is compatible with our both taking the reason, and the reason in fact being, third-person instrumental: we have reason to sacrifice only because the goddess, who is genuinely magnificent, desires it. This has implications for our deliberation. Suppose you want the sacrifice to be swift; I want the prayers to be chanted slowly. Suppose you want to sacrifice at dawn; I want to sleep in. When we reach a conclusion about how and when to sacrifice, neither of us need assume that each other's preferences have genuine normative standing in Korsgaard's sense. We might resolve our disagreements simply by pointing to what the goddess wants. If the goddess wants the sacrifice performed at dawn, then that is what we have decisive reason to do. If so, I do not respect your preferences or reasons for their own sake. I take them into account only insofar as acting on them is likely to satisfy the desires and purposes of the goddess. You treat my reasons the same way. If this is right, we can deliberate solely on the basis of reasons that are agent-neutral but that we also take to be third-person instrumental.

Suppose it is true that we can deliberate merely on the basis of reasons that are both agent-neutral and third-person instrumental. If it is, then premise (4) of Korsgaard's argument is false: it is not the case that deliberating together requires that each agent must respect the other's reasons "for their own sake," treating the other's reasons as "having authentic normative standing" simply because they are her reasons (pp. 201–02). Perhaps more importantly we can also see that the fact that the reasons we deliberate with are agent-neutral does not imply that we must treat each other's humanity as an end in itself. After all, we may both believe that we have an agent-neutral reason to sacrifice to the goddess. But we can also treat each other's humanity—each other's power to rationally deliberate—as having only instrumental worth. It is of worth only insofar as it can be employed to appease the goddess, who alone has non-instrumental worth. Premise (5) of the argument, it appears, is also false.

The second point we wish to make is this: suppose we were to grant that Korsgaard's argument is sound. (This would require us to accept premise (3), which seems to us problematic, as there might be reasons that apply to groups of limited number.) The argument implies, however, only that agents who deliberate together must treat the humanity of those with whom they deliberate as an end in itself. It has no implications whatsoever for how we ought to treat the humanity of those with whom we do not deliberate. It is consistent, then, with our holding that there are some humans who are beneath deliberating with, say, because of their ethnicity, religious convictions, sexual orientation, or the like. Indeed, one way in which people often try to humiliate others is by refusing to engage and deliberate with them. If this is so, Korsgaard's

argument does not lay the empty formalism charge to rest. This charge, after all, is not that the categorical imperative fails to yield the proper ethical conclusions in some cases. The charge is rather that it fails to yield them in a sufficiently wide range of cases. Since Korsgaard's argument has such modest implications, the empty formalism charge stands. The argument simply fails to have implications for some of the cases we care about most.

It might be replied that this last objection fails to appreciate the full import of Korsgaard's argument. For the only sort of justification, it might be said, for treating the humanity of *only* those with whom one deliberates as an end in itself would be agent-relative or private. However, there are, Korsgaard argues, no agent-relative or private reasons (p. 204). When combined with premise (3), this implies that all reasons are agent-neutral. If all reasons are agent-neutral, it might be said that (6) implies:

- (7) Necessarily, agents must treat humanity, wherever they might encounter it, as an end in itself.

Let's grant, for the moment, Korsgaard's claim that there are only agent-neutral or public reasons, which rests on a highly controversial argument (The argument is offered in ch. 9). The problem is that, even if all reasons were agent-neutral, this fact alone would not ensure that (6) implies (7). For recall that an agent-neutral reason is simply a reason that makes no reference to the agent herself. As usually understood, such a reason implies that if there is a reason for me to bring about some end, then, necessarily, there is also a reason for you to bring about that same end. Consider, now, this putative agent-neutral reason, which one might hold is implied by premise (5):

Necessarily, every agent has reason to treat only the humanity of co-deliberators as an end in itself.

Combining this claim with (6) does not imply (7). What this helps us see is that, if (6) is to imply (7), we need to identify a claim of the right sort. Given the structure of Korsgaard's argument, we could derive (7) from:

Necessarily, every agent whatsoever co-deliberates with every other agent.

But this claim is plainly false. Granted, it might be true that every other agent is one with whom one *could* co-deliberate. But even if this were true, it would establish only that:

Possibly, agents must treat humanity, wherever they might encounter it, as an end in itself.

But that is a far weaker conclusion than (7). It needn't imply that you or I actually have any reason to treat humanity, wherever we encounter it, as an end in itself. There might, of course, be a defensible claim that ensures that (6) implies (7). But that claim is not that all reasons are agent-neutral. And so far as we can tell, *Self-Constitution* does not supply it.

Perhaps there is a reason for this. For it may be that Korsgaard does not intend for the argument formulated above to put the empty formalism charge to rest. Perhaps it is supposed to be only one piece of a more elaborate case for a Kantian normative theory. If this is so, however, then there remains a challenge to furnish the remainder of the argument. That challenge, however, strikes us as daunting. The history of moral philosophy does not inspire confidence that it can be met.⁴

⁴ Thanks to Louis deRosset for comments on an earlier draft of this review.



Book Reviews

Paul Blackledge and Kelvin Knight (eds)

Virtue and Politics: Alasdair MacIntyre's Revolutionary Aristotelianism (Notre Dame: Notre Dame University Press, 2011) 384 pp. ISBN: 0-268-02225-9. Paperback: £34.95.

What is the best way to interpret Alasdair MacIntyre's political philosophy? Since the publication of *After Virtue*, MacIntyre has often been read as a nostalgic communitarian who has moved on from the Marxist politics of his youth. Recently, however, this interpretation has been questioned by a group of scholars who see an enduring radicalism in MacIntyre's mature thought. This new book is billed as a culmination of their project. It brings together a mix of philosophers, political theorists, sociologists, and historians – as well as MacIntyre himself, who contributes an introduction and reply to his critics – to reassess MacIntyre's complex political philosophy, and consider whether 'revolutionary Aristotelianism' is a more fitting interpretation of it.

The idea of 'revolutionary Aristotelianism' leads many of the contributors to consider the continuity between MacIntyre's youthful Marxism and mature Aristotelianism. According to the editors, 'MacIntyre's turn to Aristotle...is best understood not as a conservative rejection of modernity but as an attempt to deepen insights inherited from Marx's critique of capitalism' (p. 2). On this matter there is little consensus. For whilst many of the commentators agree with the editors that MacIntyre's mature thought can be usefully described in this way, others suggest that whilst MacIntyre cannot not be read as a nostalgic communitarian, he cannot be read as a revolutionary either; revolutionary Aristotelianism, as one commentator puts it, is an 'unhappy mix' (p. 79).

Appropriately the volume starts by reprinting Kelvin Knight's 'Revolutionary Aristotelianism,' the essay that first challenged accusations of conservatism back in 1996. What is revolutionary about MacIntyre's Aristotelianism? According to Knight, the answer has to do with MacIntyre's concept of a 'practice.' A practice is a cooperative human activity, engagement in which enables the discovery of goods that are 'internal' to that practice, as well as the development of practical skills and moral qualities or 'virtues.' The critical thrust of this position comes from MacIntyre's belief that practices are at odds with the central features of modern society, especially the capitalist economy.