

Two Views of Kantian Practical Reason

Amy Lara, Kansas State University

Abstract

One of the best recent attempts to provide a Kantian grounding for morality is Christine Korsgaard's *The Sources of Normativity* (1996). Her argument is powerful, but there is a deep ambiguity within it between two conflicting views of practical reason. One of these views involves a picture of the rational agent as standing outside all of her contingent practical identities and seeing the normative force of the moral law. The other view sees a rational agent as one who has many substantive practical identities that are already informed by the moral law, so the normativity of that law can be seen from within one's contingent identities. I argue that the former view is motivated by a prejudice about what practical rationality requires, a prejudice that is popular among ethical theorists in general. I also argue that the latter view is superior; it can better support the Kantian claim that the categorical imperative is intrinsically motivating, and it is less susceptible to the familiar charges that Kantian ethics is alienating.

Two Views of Kantian Practical Reason

Several pressing issues confront those who want to develop a neo-Kantian ethical theory. Common causes for concern are the charges that Kant’s theory is inherently sexist, insensitive to cultural diversity, speciesist, and alienating. I think most ethical theorists agree that if a neo-Kantian ethical theory is to be viable, it should not be inherently sexist or racist, and that it would be preferable that the theory see us as having some direct duties to animals, and not just the indirect ones Kant speaks of. In fact, neo-Kantians have made quite a bit of progress on answering these charges. There is one issue, however, that I think neo-Kantians, and ethical theorists in general, do not agree on. That issue is the question of how to ground the categorical imperative itself, and how to account for the way it is supposed to guide action. In other words: what kind of picture of practical reason will our neo-Kantian ethical theory employ?

Neo-Kantians have spent some time exploring this issue, but progress has been hampered by a certain kind of prejudice about practical reason, what John McDowell (1996) calls the “deductivistic prejudice.” The prejudice is that if we are to think of an agent as acting on a universal principle, such as the categorical imperative, we must think of the agent as applying that principle to separate facts about the situation and forming a practical syllogism modeled on deductive syllogisms. In other words, the prejudice is that if we are going to see the categorical imperative as the grounding principle of practical reason, we have to see it as separable from the particular facts an agent sees. On this kind of view, we would picture the virtuous agent as one who has a variety of particular commitments, like anybody else does, and who sees the value of many things, but who takes the *further* step of applying the categorical imperative to all her maxims and acting in accordance with the results.

This picture causes many problems for neo-Kantians, not the least of which is explaining how the categorical imperative can be intrinsically motivating. Whenever any moral theory pictures the principles of morality as graspable from a standpoint completely divorced from our actual practices, it becomes very easy for a non-cognitivist to step in and argue that these principles are not intrinsically action-guiding. After all, the moral principles can be grasped by the most immoral criminal just as well as by the most virtuous agent, but that criminal seems to be completely unmoved by them. Thus, these moral principles seem to give a reason to be moral only to the person who also has a desire to follow them, and that desire is not itself necessary. Clearly, a neo-Kantian will want to resist such a view, since one of the central claims of Kantian ethics is that the categorical imperative gives a reason to be moral to every rational being, regardless of his or her desires.

Because this is a central claim of Kantian ethics, as well as a hot topic of dispute in ethical theory, ethical theorists in general should be interested in examining how neo-Kantians attempt to support the claim. If McDowell is right about the deductivistic prejudice and the problems it creates, then we would expect to see neo-Kantians resisting that prejudice in order to argue that the categorical imperative is intrinsically motivating. Yet, can they do this without undermining the universality and inescapability of the categorical imperative? If so, that would open up an interesting new way of grounding morality for Kantians and non-Kantians alike. On the other hand, if neo-Kantians are able to maintain the deductivistic picture and still show that the categorical imperative is intrinsically normative, that would also be interesting. Arguments made by McDowell and others that we ought to move toward more Aristotelian virtue theories would lose much of their force.

An excellent place to start in examining this issue is by looking at Christine Korsgaard’s *Sources of Normativity* (1996). This very influential book, comprised of Korsgaard’s 1992 Tanner Lectures, provides a beautiful example of a neo-Kantian attempt to ground morality. The resulting view is clearly rooted in Kant, but also modifies Kant’s view in many ways to make it stronger. In particular, Korsgaard takes herself to be responding to charges of alienation that have been leveled against Kant’s ethics (1996, p. 239), and she also provides a strong argument that we have direct duties to non-human animals (1996, pp. 149-157).

This book is such a good point of entry into these issues because it tackles head-on the problem of showing how the categorical imperative can be intrinsically motivating. As I will argue in this paper, Korsgaard herself clearly struggles with the deductivistic prejudice throughout her arguments. On the one hand, she often gives in to the temptation to ground the universality of the categorical imperative in a standpoint completely distanced from the particular reasons for action that moral agents see. On the other hand, she senses that this strategy will undermine the categorical imperative’s motivational force, and she ends up building the CI itself into all of our particular reasons for action. We end up with a view that relies on two incompatible pictures of practical reason.

I will start my argument with an overview of Korsgaard’s project in *Sources of Normativity*. I will then examine two of her particular arguments, and show that these two arguments rest on different pictures of practical reason. I will then argue that one of these pictures is motivated by the deductivistic prejudice, and that this creates serious problems for the resulting view. Finally, I will argue that the non-deductivist picture is the superior view and can support the view Korsgaard wants.

Overview of Korsgaard’s Argument

In *The Sources of Normativity*, Korsgaard’s ambitious project is to ground the normativity of moral requirements. She takes a largely Kantian approach to this project, but makes some arguments that are quite different from Kant’s, and the resulting view departs from traditional Kantian ethics in important ways.

Korsgaard begins her argument by examining the capacity for reflection that we have as rational beings. She says, “our capacity to turn our attention on to our own mental activities is also a capacity to distance ourselves from them, and to call them into question” (1996, p. 93). This capacity generates a problem for us: once we have distanced ourselves from our own “mental activities,” especially from our own desires, we need a way to decide how to move forward, a way to decide which desires we will act on, or whether we’ll act on our desires at all. In other words, we now need a reason for action, rather than just a desire to act on. Korsgaard calls this “the normative problem.”

At this point, Korsgaard, following Kant, brings in the notion of “the alien” (1996, p. 97). The capacity to step back from our desires and guide our actions by principles instead, is autonomy, or freedom from alien influences. To act on a principle of reason rather than directly on a desire is to act autonomously rather than heteronomously. But how shall the will choose? Once we have a distance from our desires, we have to figure out how to decide which desires to endorse, as the desires themselves are no longer intrinsically normative. Korsgaard’s answer is this:

The problem faced by the free will is this: the will must have a law, but because the will is free, it must be its own law. And nothing determines what that law

must be. *All that it has to be is a law.* Now consider the content of the categorical imperative, as represented by the Formula of Universal Law. The categorical imperative merely tells us to choose a law. Its only constraint on our choice is that it has the form of a law. And nothing determines what the law must be. *All that it has to be is a law.* Therefore, the categorical imperative is the law of a free will (1996, p. 98).

Here Korsgaard gives us an argument that is quite close to Kant’s own argument for the Universal Law formulation of the Categorical Imperative. The idea is that the Categorical Imperative, at least in this formulation, is a constitutive principle of autonomous willing. It describes what a free will must do in order to be a free will: a free will must act on maxims that it can regard as having the form of law (1996, p. 98).

From this point, though, Korsgaard’s argument departs radically from Kant’s. Korsgaard contends that the Kantian argument given so far only establishes that the Universal Law formulation of the Categorical Imperative is constitutive of autonomous willing. It only shows that as free, rational beings we must act on laws. In order to show that we must act on *moral* laws, however, we will need a further argument. As Korsgaard puts it:

Any law is universal, but the argument I just gave doesn’t settle the question of the *domain* over which the law of the free will must range. And there are various possibilities here. If the law is the law of acting on the desire of the moment, then the agent will treat each desire as a reason, and her conduct will be that of a wanton. . . . It is only if the law ranges over every rational being that the resulting law will be the moral law (1996, p. 99).

So we know at this point that we must act on universal laws, but we don't yet have the moral law, which is Korsgaard's name for the Kingdom of Ends formulation of the Categorical Imperative. The moral law says that we should “act only on maxims that all rational beings could agree to act on together in a workable cooperative system” (1996, p. 99).

How will we bridge the gap from the Universal Law formulation of the Categorical Imperative (which Korsgaard calls simply “the categorical imperative”) to the moral law? This is where Korsgaard brings in the notion of a practical identity, a notion that is not used by Kant. A practical identity is “a description under which you value yourself, a description under which you find your life to be worth living and your actions to be worth undertaking” (1996, p. 101). The need for such an identity arises from reflection itself. Reflection forces us to choose which desires to endorse, and the principle by which we choose expresses our conception of ourselves. This is unavoidable because reflection requires us to choose independently of alien influences; the choice made from this position, then, expresses what we take to be alien and what we take to be self. Reflection forces us to have a conception of ourselves.

Korsgaard uses this notion of a practical identity to bridge the gap between the categorical imperative and the moral law. Her argument follows this outline: (1) There are some ways we must think of our practical identities; namely, we must identify as human beings, or value our humanity. (2) Valuing our humanity necessarily involves valuing humanity in general. (3) So, we must identify as fellow humans with others; that is, we must have a moral identity that generates the moral law as an obligation for us.

Once this argument is made we have an ethical theory that is roughly Kantian, but that varies from Kant in important ways. On Korsgaard's view the normativity of all of our reasons for action is grounded in our practical identities, and we can have many practical identities at

once. The normativity of moral reasons for action, however, is grounded in our identity as rational human beings. This identity has priority over our other identities, and all of our practical identities are partly grounded in that human identity. All of this is pretty distant from Kant. The Kantian part however is seen in the claim that our identity as rational human beings is an identity as autonomous agents who need reasons for action and who don't just act directly on their desires.

I want to turn now to the particular arguments Korsgaard gives for the first two premises I outlined above. The first premise is that we must identify as humans and value our own humanity. The second premise is that valuing our own humanity necessarily involves valuing humanity in general. The argument for the first premise relies on a very interesting picture of practical reason, one motivated by the “deductivistic prejudice” I mentioned earlier. However, the argument for the second premise relies on a quite different picture of practical reason, one that is not compatible with the first picture. Once I bring out these two pictures I'll argue that the second is the stronger one, and that the first should be jettisoned.

The Argument for Premise One

The argument for premise one needs to show that we must value our own humanity. Remember that at this point in the argument we've already seen that reflection requires us to act on a universal law, and to have some kind of conception of ourselves. What Korsgaard needs to show now is that we must conceive of ourselves in a certain way, as rational beings who act on

laws, and that this self-conception requires valuing our own humanity.¹ Notice, however, that there are two different ways we might make this argument. What we have so far is a transcendental argument that choosing in accordance with universal law is constitutive of free willing. We could make the purely logical point that the activity of choosing in accordance with universal law is itself a valuing of humanity, and thus that every choice we make already involves a valuing of humanity. This would mean that having a conception of ourselves as human would itself be constitutive of the process of making decisions. We could then say that that self-conception is inherent in all of our practical identities, and is not a separate identity as human. On the other hand, we can approach the argument psychologically. We can think of the agent herself as seeing that she needs to have a practical identity and wondering whether there are any identities that are necessary. She could examine and question each identity until she arrives at one that is necessary, and that would be the separate human identity. Korsgaard takes this latter approach, at least in this part of her argument.

The psychologizing of Kant’s argument shows up early. Korsgaard frames “the normative problem” itself in terms of an individual agent’s thought process, and she explicitly contrasts Kant’s approach to this problem with the neo-Humean approach that theorists like Bernard Williams take. She writes:

[As] each impulse to action presents itself to us, we should subject it to the test of reflection, to see whether it really is a *reason* to act. . . . Hume and Williams see the test of reflective endorsement as a philosophical exercise, used to establish the normativity of our moral dispositions and sentiments. But according to Kant, it is

¹ Following Korsgaard and Kant, I use “humanity” simply to refer to us as rational animals, not to refer to us as members of a particular species.

not merely that. The test of reflective endorsement is the test used by actual moral agents to establish the normativity of all their particular motives and inclinations. So the reflective endorsement test is not merely a way of justifying morality. *It is morality itself* (1996, p. 89).

Clearly, then, Korsgaard is conceiving of the moral agent as constantly engaging in a process of reflection on her desires and interests, so that the categorical imperative is not just a description of that process that we might give after the fact, but is supposed to be an implicit or explicit guide that we follow as we make choices.

This psychologizing of the argument is in fact what opens the gap between the categorical imperative and the moral law. We are to imagine an agent examining each desire, trying to decide which to endorse. She has no unquestionable principle of choice to use, except that her choice must have the form of a law. But which law? It looks as though she could set any law at all for herself, as long as it is universal in form. Further argument is needed to establish the moral law as the law to adopt. And it's not just that we as philosophers need that further argument. The agent herself requires a further process of thought in order to see what law she should act on.

Now that the psychological problem is set up, Korsgaard is ready to provide the solution. We know that the agent needs a principle, which will express her self-conception, and thus she needs a practical identity. So now the problem is which identity to have; or, given that she already has some identities in place, which should she endorse? Korsgaard's argument follows this outline: Most of our particular practical identities are contingent and relative; we find ourselves born into a particular family, citizens of a particular country, lovers of particular people. Each of these identities gives rise to reasons for action, and these are the reasons upon

which we usually act. However, because these identities are contingent, we can give them up if we decide not to identify with them anymore. What we cannot do, though, is give up on having a practical identity at all. We must have some practical identity or other in order to have any reason to act at all. This is a reason for conforming to our particular practical identities, but it “is not a reason that *springs from* one of those particular practical identities. It is a reason that springs from your humanity itself, from your identity simply as *a human being*, a reflective animal who needs reasons to act and live” (1996, 121). There is, then, a practical identity that is not contingent: our identity as human beings, as beings who are reflective and need reasons to act.

Korsgaard conceives of this human identity as similar in structure to our other practical identities, but as having a special status because of its inescapability. As agents with contingent practical identities, we do have a noncontingent reason to act in accordance with those identities, but this is “a reason you have only if you treat your humanity as a practical, normative, form of identity, that is, if you value yourself as a human being. But to value yourself as a human being is to have a moral identity...” (1996, p. 121). Korsgaard says this moral identity is, in a sense, “just like any other form of practical identity” (1996, p. 129). The difference is that the moral identity helps ground the normativity of our other identities, and also the “moral identity exerts a kind of governing role over the other kinds” (1996, p. 130).

At this point, Korsgaard hasn’t actually made the full argument that a human identity is a moral identity, and I will discuss that argument in the next section. For now, I just want to focus on the particular picture of practical reason that is being developed in this part of the argument. Notice, first, that the argument implies that in order to be practically rational, an agent’s reasons for action must have a certain hierarchical structure. The reason-giving force of the particular

reasons an agent has is grounded in the primary reason for action that the agent can grasp from the standpoint of reason: namely, that one must have a practical identity. This is the reason that stands behind and legitimates the alleged normativity of our other reasons for action. Our particular identities, then, don't generate full reasons for action on their own. The force of the reasons they generate is derivative from the force of reasons generated by our identity as humans. Humanity is the source of normativity.

Implicit in this hierarchical view is the idea that there are two separate standpoints a rational agent can take: the standpoint of any or all of her particular, contingent identities, and the standpoint of universal reason. These two standpoints are not only logically separate, but psychologically separate. The standpoint of reason is the standpoint of distance from one's desires and particular identities, the standpoint from which one can question the normativity of those things. This means that, from the standpoint of reason, one's particular identities are to a certain extent *alien*. The agent needs to evaluate their purported normativity, and then choose freely to adopt them if they meet the standards of rational endorsement. This requires that the agent have some sense of a self from within the standpoint of reason, a self that is separate from these particular identities.² For this reason I will call this view the Separate Standpoint View (SSV).

Another interesting thing about this view is that it doesn't especially matter which particular identities we have when we enter the standpoint of reason. As long as we have some, we can enter that standpoint. And as long as the ones we have are morally permissible, we can

² Korsgaard is explicit about this point: “When you deliberate, it is as if there were something over and above all of your desires, something which is *you*, and which *chooses* which desire to act on” (1996, p. 100).

embrace them from that standpoint. So our particular training and desires don't play any crucial role in our ability to reach the standpoint of reason. This makes the standpoint of reason accessible to a wide range of people, presumably including amoral people. So the argument up to this point supports the idea that we must all identify as human, and be governed by the categorical imperative, regardless of our particular practical identities and training.

This view of practical reason, I contend, is motivated by the deductivistic prejudice that I mentioned earlier. McDowell characterizes the prejudice this way:

The idea . . . is that the content of a conception of the universal, doing well, is in principle available, and assessable for correctness, in abstraction from the judgments or actions, in particular circumstances, that we want to see as applications of it. . . . Correctness of application would be recognizable, in principle, from a stance that was neutral with respect to the corresponding end (1996, p. 22).

Here McDowell is criticizing an assumption that he sees as common in modern ethical theorizing: that in order for an action to be genuinely rationally guided, it would have to result from knowledge of an abstract or formal rational principle that is applied to particular cases. For example, we might think of the Kantian virtuous agent as committed to the principle that she should never act on maxims that cannot be universalized. She then runs maxims through the universalizability test and sorts them into permissible and impermissible. On this picture, even a complete outsider who has no commitment to the categorical imperative could fully understand the principle to which this virtuous agent is committed, and could evaluate the correctness of her application of it to particular maxims. To prove that this virtuous agent is indeed acting in the most rational way, we would have to show that her testing procedure is accurate and consistent,

and also show that her general principle is the rational one to follow. The critical aspect of this picture is the idea that the general principle of the virtuous agent could be assessed for correctness independently of its ability to give us the right answers when it is applied to particular maxims.

The Separate Standpoint View rests on the above picture. In arguing for premise one, Korsgaard assumes that in order for us to see that a particular universal law, such as the moral law, is binding on us, we must be able to enter a completely universal standpoint that contains no contingent content. If we can see the force of the moral law from that standpoint, then it will have the kind of inescapability that we want the moral law to have. So in Korsgaard's framework, the deductivistic prejudice about how to apply universal principles becomes a prejudice about what kinds of practical identities we must have as moral agents. The moral identity itself can only be a universal identity with no contingent content.

We haven't yet seen whether this prejudice is actually a problem, and I will discuss that shortly. Now, however, I want to look at the argument Korsgaard makes for Premise Two, that we must value humanity in general.

The Argument for Premise Two

At this point in the argument, Korsgaard takes herself to have shown that we must identify as humans and value humanity. She believes that this implies that we must value humanity in general, not just our own humanity, but she realizes that there may be skeptics about that implication. So she spends some time showing that all reasons for action implicitly involve

a valuing of humanity in general, so that the very process of making choices and identifying ourselves as humans already involves valuing the humanity of others.

To make this argument she brings in a new distinction, one that fundamentally alters the view of practical reason that she is working with. The distinction is between public and private reasons. Private reasons are reasons for action that have normative force only for an individual agent. Public reasons have normative force for all agents (1996, p. 133). Korsgaard complains that ethicists, including Kantians, have often assumed that public reasons must be reduced to private reasons, and that the only non-controversial private reasons we have are self-interested ones. For example, a Hobbesian might argue that we all have reason to be moral because we each have a private reason to act for our own self-preservation, and morality is the best way to do that. Similarly, a Kantian might argue that we each have a private reason to value our own humanity, and this implies that we all have reason to value each other's humanity.

Korsgaard acknowledges that Kantian arguments from private to public reasons won't work. Valuing my own humanity does not imply that I must value yours; it only forces me to acknowledge that your humanity has value for *you*. This means that an argument from the categorical imperative to the moral law won't work if it argues from the CI to the idea that I have a private reason to value my own humanity, for then there will be no way to argue from this to the moral law itself.

Korsgaard's solution is to argue that there are no such things as private reasons to begin with. Her thesis is that “to act on a reason is already, essentially, to act on a consideration whose normative force may be shared with others” (1996, p. 136). All reasons for action are public. This is because reasons involve normativity, and all normativity is relational:

To say that R is a reason for A is to say that one should do A because of R; and this requires two, a legislator to lay it down, and a citizen to obey . . . [so] it takes two to make a reason. And here the two are the two elements of reflective consciousness, the thinking self and the acting self (1996, pp. 137-138).

Another way to put this is that reasons conceptually involve a relationship of authority. One member of the relation has the authority to obligate the other member to do something. Only with this authority in place is the other member given a *reason* to do something. It then becomes the case that that member *should* do the thing in question.

But why do we say that this is a relation between two different entities? Why can't there be private reasons where one entity obligates itself to do something? The reason is that no sense can be made of obligation when there is no relation of authority between two entities. If obligation were completely private, given by one entity to itself, then there wouldn't be any way for that entity to fail to meet the obligation. Whatever one did would count as meeting the obligation because there's no independent standard of right and wrong.

Korsgaard appeals to Wittgenstein's private language argument to make this point. Linguistic meaning is a normative concept, just as reasons for action are normative. Because meaning is a normative concept, it must be the case that one can go wrong in following the norm. But in a completely private language one cannot go wrong. As Wittgenstein puts it, "whatever is going to seem right to me is right. And that only means that here we cannot talk about 'right'" (Wittgenstein 1953, p. 92; Korsgaard 1996, p. 137). A criterion of correctness, then, must be inherently public.

Does this mean, then, that a person can never obligate herself to do something? Not if we think of the person as having two parts that stand in a relation of authority to each other. And it

is reflection that allows us to think of a person in this way. In reflection the agent is distanced from her particular desires, wondering which to endorse. The distanced self is what Korsgaard calls “the thinking self.” At the same time there is still the self who has these desires, these values, and this is “the acting self.”³ Reflection sets up a relation between these two selves wherein the thinking self is given the authority to endorse and withhold endorsement of the acting self’s desires; this obligates the acting self to act in certain ways.

This distinction between the thinking and acting selves would be only nominal, however, if we did not also add the idea that the thinking self is inherently public. As a reflecting being I must think of my thinking self as occupying a public space, a place from which it sees *correctly* which desires to endorse. It has to be the case that in principle others can occupy this position if they meet the same criterion for occupying it that my thinking self meets. And this criterion is simply being able to achieve distance from one’s desires and act on principles. In other words, humanity is the criterion.

To bridge the gap between the categorical imperative and the moral law, then, Korsgaard needs to show that in reflection, we must acknowledge a public reason (a reason shareable with others) to value humanity as an end, not merely a private reason to value our own humanity. Korsgaard makes that argument with the following type of example: suppose that a stranger, Jones, is tormenting you, and you tell him to stop by asking how he would like it if someone tormented him in this way. Jones realizes that he would resent being tormented and that the tormenter would have an obligation to stop that would be rooted in Jones’s *own* objection to the tormenting. Jones’s humanity directly generates an obligation for the tormenter. But if Jones’s humanity generates such an obligation, so does your humanity, and Jones is equally obligated to

³ The distinction between the thinking self and the acting self is explained on page 104 of (1996).

stop tormenting you. Jones’s reflection does not merely generate an obligation for Jones to value his own humanity, but an equal obligation to value yours (1996, pp. 142-143).

This argument for the bindingness of the moral law may not be convincing yet. This is because there is a gap between the proposition that reasons are not completely private, which Korsgaard argues for, and the proposition that reasons are universal and are only binding if they are acceptable in principle to all humans, which is ultimately what the argument must demonstrate. For example, why couldn’t there be reasons for action that I view as shareable only with other members of my family or race?

The temptation here is to fall back on the Separate Standpoint View. Why not say that in the Jones example what you are doing is asking Jones to step into the standpoint of reason, to question his own motivations to torment you? Once he is in the standpoint of reason, Jones can see that his desire to torment you couldn’t have any necessary reason-giving force unless that force were grounded in Jones’s own identity as a being who needs reasons, and not just desires, to act on. Thus, to take the desire as reason-giving, Jones would need to value his own humanity. But, then, aren’t you human too? Isn’t it inconsistent for Jones to value his own humanity, his status as a being who needs reasons, while not valuing that same exact status in you?

But Korsgaard resists this way of arguing. She thinks that if reflection only commits Jones to valuing his own humanity, and then further argument is required to force Jones to value others’ humanity as well, the argument will fail. No argument from logical consistency can force Jones to value your humanity *himself*; it can only force him to see that you must value your own humanity just as he values his. What Korsgaard says instead is:

There is an appeal to consistency in this argument; it is meant to remind [the tormentor] of what the value of humanity requires. But it is not what makes [him] take [your] reasons into account, or bridges the gap between [his] reasons and [yours], for there is no gap to bridge (1996, p. 143).

In other words, when you ask Jones how he would like it if you tormented him, you are simply inviting him to change places with you. But that invitation only works insofar as Jones can already see himself as one person among others. For the invitation to fail, Jones would have to see himself as uniquely human, and see you as not fully human. And Korsgaard says, “the argument never really fails in *that* way. For it to fail in that way, [he] would have to hear your words as mere noise, not as intelligible speech. . . . In hearing your words as *words*, [he] acknowledge[s] that you are *someone*” (1996, p. 143). Thus, reasons cannot just be shareable with one’s race or family. Seeing another person just as a language speaker already puts him or her into the realm of humans with whom one shares reasons.

If this argument works, it establishes a surprisingly strong conclusion. The conclusion is that all reasons for action have normative force for all human beings, and that we implicitly acknowledge that every time we try to act on reasons and not just on desires. Furthermore, we implicitly acknowledge others as full human beings every time we speak to them and listen to them. This doesn’t mean we will always, or even mostly, live up to this implicit valuing of humanity, but just by being in the business of exchanging reasons with each other we are already committed to trying to value humanity.

Yet we’ve seen that this argument that all reasons are public is in tension with the Separate Standpoint View, which is the view of practical reason that Korsgaard seemed to be working with earlier. In fact, I believe that adding the notion of public reasons to her view

radically changes the picture of practical reason Korsgaard is relying on. If we accept the thesis that all reasons for action are public, and practical identities are defined as self-conceptions that generate reasons for action, then publicity is already built into our practical identities. The reasons for action that spring from our identities are public, so they are shareable with others. Thus, acting on them involves the notion of a lawgiver and a citizen, of publicly shareable normative authority.

This means that the standpoint of reason is implicit in the standpoints of our particular identities, and need not be seen as separable from them. In generating public reasons, a practical identity already involves a distancing from one's own desires and a questioning of which to endorse. This questioning and striving to act on a shareable reason implies that one is identified as a human already. Thus, we can see “the standpoint of reason” as a way of talking about an aspect of the thought process that we engage in from *within* our practical identities, not as an actual separate standpoint that we enter when we leave the standpoint of our practical identities. I will call this view the Integrated Standpoint View (ISV).

Since, on this view, the standpoint of reason is an integral part of our particular identities, it does matter what particular identity we start with. We cannot enter the standpoint of reason from just any practical identity. Or, rather, not just any conception of ourselves will count as a genuine practical identity in the first place. A practical identity must generate reasons for action, and reasons implicitly involve a valuing of humanity, so an identity that is completely opposed to the value of humanity or to the notion of publicity will not be a genuine practical identity. The sociopathic identity, for example, if the sociopath is conceived as a complete wanton, would count as a mere cluster of desires or perhaps a temperament, rather than a coherent conception of

oneself and of what matters that generates reasons for action.⁴ Other identities that partially acknowledge the value of humanity will be inherently unstable – they will generate reasons whose normative force implies a valuing of humanity, but some of those reasons will also work against that value.

Another interesting feature of the Integrated Standpoint View is that it can hold the thesis that practical reason is a practice that we must be trained into through the process of developing a substantive practical identity.⁵ We acquire reflective distance and an ability to see the value of humanity by learning how to be a good parent, a good citizen, an artist or athlete or short-order cook. This thesis does not follow from this view of practical reason, since one could hold that the standpoint of reason is *both* implied in our practical identities *and* a separate standpoint that one can enter spontaneously, with no substantive training. However, the thesis that practical

⁴ This is not to say that such a person is actually possible. Korsgaard argues elsewhere (1997) that even instrumentally rational action requires autonomy. If that argument works, then a person who is truly instrumentally rational but completely unmoved by the categorical imperative is actually logically impossible.

⁵ Incorporating this thesis would bring the resulting view much closer to McDowell's. In McDowell's picture, morality cannot be codified into a set of principles that are graspable independently of the virtuous person's form of life. Only through proper upbringing does a person come to see moral facts properly, and her seeing particular facts properly is just what it is for her to grasp moral principles. See especially (1998b) and (1998c). The advantage of this kind of view is that it's easier to defend the thesis that moral knowledge is intrinsically motivating. On this kind of view, we *can't* conceive of a person who grasps the same facts as the virtuous person, but who is unmoved by them.

reason must be the result of substantive training is a natural companion to the Integrated Standpoint View.

This means that the Integrated Standpoint View opens up a way for us to picture practical reason without giving in to the deductivistic prejudice. On the Integrated view we can say that the moral law, the principle that we should act only on maxims that can in principle be agreed to by all other rational beings, is built into our particular reasons for action. The very practice of forming reasons for action is inherently governed by the moral law. This opens the way to say that the moral law is not fully intelligible independently of the particular reasons for action that we see. If we did not already see the point of our various reason-giving activities, we would not be able to understand the moral law as a requirement on our actions.

Comparing the Two Views on the Question of Moral Motivation

We've seen so far that there is a tension in Korsgaard's argument between two general pictures of practical reason. The question now is whether one of the pictures is better than the other, and if so, could Korsgaard move completely to that picture and still support her argument? I want to argue that the Integrated Standpoint View is superior to the Separate Standpoint View on two issues: the issue of moral motivation and the issue of alienation. Now Korsgaard was drawn to SSV because it helped her make her argument for her first premise, so we'll need to return to that argument and see whether we can replace SSV with ISV.

Remember that premise one is that we must value (at least) our own humanity. To argue for this, Korsgaard shows that our particular identities are contingent and can be given up, but

that we cannot give up on the need to have some particular identity or other. Thus, we must identify as beings who need identities, as beings who need reasons to act.

This argument encourages us to picture the agent as stepping back from her practical identities:

Being human, we may at any point come to question the normativity of one or another of our practical identities. . . . Most of the ways in which we identify ourselves are contingent upon our particular circumstances, or relative to the social worlds in which we live. How can we be bound by obligations which spring from conceptions of our identity which are not in themselves necessary? (1996, p. 129).

Once the agent has stepped back this way she finds that there is an identity that “stands behind” her particular identities: her human identity. She can question the normativity of this identity, but must ultimately come to embrace it because she *is* human.

At least in this argument, Korsgaard means for this identification as human to be a concrete, psychological process that moral agents go through. As we saw earlier, she says that this identity as a human is just like our other practical identities in structure, and that this human identity *is* the moral identity. This implies that the human identity is both necessary and sufficient for grounding morality. It is necessary because we can only see the necessary value of humanity from this standpoint, and it is sufficient because this standpoint involves identifying as a being who needs reasons, and identifying in this way implies a valuing of humanity.

But there are good reasons to doubt that something as thin and easily attained as an identity simply as a being who needs reasons could be either sufficient or necessary to ground moral requirements. Because the identity is so thin, and is attainable from any starting point, it is

unclear how it can be intrinsically motivating. On the other hand, because the identity is so distanced from our other identities, and requires questioning all of our identities, it seems to make morality alienating. I will discuss the motivation problem in this section, and the alienation problem in the next.

For the Kantian, the moral agent is supposed to be moved directly by the recognition that action is morally required or forbidden. She is not supposed to require an extra desire for the end, or a desire to avoid punishment in order to act morally. On the Separate Standpoint View this is explained by the priority of the standpoint of reason. Our contingent identities propose various courses of action, but the standpoint of reason has the final say – either it finds the proposed actions permissible, required, or forbidden. If the standpoint of reason does not have this voice in the agent’s decision, her action is heteronomous. It’s possible on such a view to suppose that an agent might develop a particular identity that is so perfect that the standpoint of reason only needs to endorse it once, and then the reasons generated by that identity could be trusted from then on. However, it’s implausible to suppose there’s any substantive identity that would be guaranteed never to generate demands that conflict with morality (as long as we’re conceiving of this identity as logically separate from the standpoint of reason itself). So the moral agent will need constantly to consult her rational identity in order to check on the incentives arising from her contingent identities.

This view raises problems for moral motivation. Sooner or later the agent will experience a conflict between a reason generated by one of her particular identities and a reason generated by her rational identity. And the particular reason will not just be a random desire or whim; it will be a reason that makes sense within that identity, where the identity is something that helps her make sense of her life. At this point the rational agent is supposed to see that her

rational identity is more important than her contingent identity, that it is in fact what stands behind that contingent identity. But because the competing reason comes from an important, though contingent identity, it's very plausible to suppose that the agent will act on the competing reason instead of the reason generated by her rational identity. We can even imagine her acknowledging that her humanity is more important than her identity as a mother, friend, or customer service rep, but right now it's still more important to act from that contingent identity. It looks, then, as though her rational identity is not intrinsically motivating – the agent needs to acquire the right set of desires not only to act on that identity, but to act against her other identities when they conflict with it. The problem becomes especially acute when we remember that the rational identity can be stepped into at any time, from any starting point, simply by distancing oneself from one's contingent identities. Since we can regain that identity very easily at any future time, what's the harm in dropping that point of view whenever we really want to, planning on readopting it right afterwards? We may then feel guilt for violating it, but we can simply resolve to act rationally from then on.

Presumably Korsgaard will answer this problem by arguing that an identity as human involves more than just recognizing the bare fact that the human identity is logically prior to our other identities. To really have this identity is to be moved by it, to see its importance and the way it makes our other identities possible.⁶ But can she help herself to this argument if she's working with the Separate Standpoint View? On the Separate Standpoint View, we achieve the human identity by completely stepping back from all our other identities. This means that we

⁶ Furthermore, Korsgaard needn't show that this human identity will in fact always trump our other identities; she only needs to show that *insofar as we are rational* we will not act against this identity. She makes this point in (1997), page 236.

now have no particular reasons for action in front of us, and we have even distanced ourselves from any conception of how to choose well that is peculiar to a contingent identity. All we have now is the knowledge that we are beings who need principles, and not mere desires, on which to act. But can we even understand what that description means, much less what it requires, from this distanced standpoint?

What exactly does it mean to act on a principle as opposed to a desire? We can't specify the difference in terms of psychological facts for two reasons: first, there doesn't seem to be any consistent phenomenological difference between the two types of action. Being moved by a reason doesn't "feel" different from being moved by a desire, and simply chanting some principle to ourselves as we act certainly doesn't guarantee that we aren't acting on a mere desire. Second, even if we could specify some phenomenological difference between the two types of action, that wouldn't be a relevant kind of difference. What we want to know is what the *normatively significant* difference between them is supposed to be.

We can get a clue from Korsgaard's argument for why it's important to act on a principle rather than just on desires: "if *all* of my decisions were particular and anomalous, there would be no identifiable difference between *my acting* and *an assortment of first-order impulses being causally effective in or through my body*. And then there would *be* no self – no mind – no me – who is the one who does the act" (1996, p. 228). The difference, then, is that actions on reasons are genuine actions, while "actions" caused by desires are mere behaviors. This is because reasons for action are self-given and free, while desires are alien causes (1996, p. 97). Thus, the reason/desire distinction, *as a normatively important distinction*, is parasitic on the self/alien distinction.

Now the question is whether the self/alien distinction can be fully grasped simply from the point of view of our human identity. If so, the self/alien distinction would need to explain the reason/desire distinction in a way that can guide us by providing reasons for action. The self/alien distinction explains the reason/desire distinction in this way: reasons are those things given by the will to itself, while desires are alien causes that come from outside the will. This is important because we strive to act freely, not just be pushed about by causes. This explanation does tell us why we should care about the distinction between reasons and desires, but it doesn't actually get us any closer to being able to distinguish between reasons and desires. What makes a motivation internal to the will rather than external to it? Again, the difference cannot just be phenomenological. It wouldn't help if we just said “I will that ...” before each maxim we formulated. Neither can we introspect and get a special feeling of a motivation coming from “us” as opposed to coming from “outside.” And even if there were a special feeling accompanying all of our “self-given” motivations, it would be normatively irrelevant. It wouldn't tell us what made that class of motivations self-given in the relevant sense.

I think it's telling that whenever Korsgaard feels the need to explain the importance of acting on reasons, and not just on desires, she asks us to imagine ourselves making a particular decision:

From a third-person point of view, outside of the deliberative standpoint, it may look as if what happens when someone makes a choice is that the strongest of his conflicting desires wins. But that isn't the way it is *for you* when you deliberate (1996, p. 100).

And, again, when she's comparing the subjectivity of colors to the subjectivity of willing:

If you think colours are unreal, go and look at a painting by Bellini or Olitski, and you will change your mind. If you think reasons and values are unreal, go and make a choice, and you will change your mind (1996, p. 125).

Korsgaard sees the need to place us in the first-personal perspective of actually deciding how to act in order to show us the importance of reflection and of acting on reasons. It is only from the perspective of actual deliberations that we have a concrete sense of self as a chooser, and of what counts as a reason as opposed to a mere desire.

This gives us a way to ground the normativity of the self/alien distinction: ground it in the particular practical identities that we have. Our identities provide us with a way of distinguishing external desires from self-given motivations; that is in fact how those identities work to give us a self-concept. For example, an athlete learns which feelings of fatigue and pain should be worked through, and which should be taken as reasons for resting. A mother hates her children's pain, but this is seen as an extraneous desire to be fought against if it interferes with doing what's best for the child. To understand the difference between an extraneous desire and a genuine reason for action, and to understand why it's important to act on the latter and not the former, we need to see that distinction in action, to see it in the context of actual identities we have.

So the human identity can't just be one identity among others. If we try to spell it out as a separate identity, its normativity fades away. When we ask the question of whether to act on a certain desire, we don't even know what would constitute an acceptable type of answer to that question except from the standpoint of our particular identities. So only the Integrated Standpoint View is going to be plausible for explaining the normativity of the categorical imperative.

Yet, can the Integrated Standpoint View still support the universality and inescapability of the CI? If we can only see the importance of humanity, of the capacity for choice itself, from the standpoints of our contingent identities, then doesn't that make the categorical imperative itself contingent? The answer is that there is one way in which our practical identities are not contingent: they are all ways of forming a self that is separate from the particular desires that move us. If they did not provide principles for action, they would not be true practical identities. And, as Korsgaard has argued, the categorical imperative is a constitutive principle of the practice of acting on principles rather than desires. So the categorical imperative is inescapable, and governs all practical identities, because that's what it is for something to be a practical identity. However, lest this appear to be a reiteration of the Separate Standpoint View, we need to remember that the fact that a practical identity is the kind of thing that is governed by the CI is not itself graspable from a neutral perspective outside of our practical identities. To understand the point of the categorical imperative, we actually have to imagine ourselves confronting some concrete reasons for action.

A further consequence of moving completely to ISV is that it closes some of the gap between the CI and the moral law, bringing Korsgaard's view back in line with Kant's texts. Recall that in Korsgaard's argument for premise one, the gap between the CI and the moral law opened up because the agent was pictured as being able to grasp the necessity for acting on some law or other, without yet seeing that she needed to act on the moral law. Further argument was required to establish the objective normativity of the moral law. But if we start the argument over with ISV clearly in mind, the picture is very different. For an agent to see the need to act on universal law is for that agent already to be in the business of acting for reasons, and thus for the agent already to be valuing humanity as an end. Understanding the CI as a normative

requirement already requires seeing the point of acting on reasons, of acting in ways we can justify to others. But, conversely, acting in a justifiable way already implicitly involves the idea of acting on maxims that have a lawful form, that are not mere individual impulses. So while ISV still allows us to draw a *conceptual* distinction between the CI and the moral law, we can retain Kant’s own claim that the CI and the moral law mutually imply each other (Kant 1997, p. 43; 4:436).

Comparing the Two Views on the Question of Alienation

There’s another reason to prefer the Integrated Standpoint View: it gives us an ethical theory that is much less susceptible to charges of alienation. The Separate Standpoint View requires us to picture the agent as taking various standpoints on her desires – the standpoints of her identities as well as the standpoint of reason. Even when her particular identities give rise to reasons that her human identity can endorse, the agent still needs to consult that human identity to be sure. This leads to what Bernard Williams has called “one thought too many.” Adapting his example, when a man chooses to save his wife first after a boating accident, his choice is motivated by his identity as this woman’s husband. As her husband he will of course look to her first, before considering other victims who are strangers. Yet this husband-identity is contingent and partial, and can in principle give rise to reasons that conflict with morality. So the man also must act on his human identity, which presumably will say that it’s okay for people in general to save their own family members first in this kind of situation. This is the “one thought too many.”⁷

⁷ For the original discussion, see Williams (1981b), pages 17-19.

Now it would be unfair to say that the Separate Standpoint View requires an agent to step into each standpoint separately in the moment of deliberation. This view need not hold that a person enter the standpoint of reason each time she deliberates, or even that a person need deliberate each time she acts. But the view does require that at some point the moral agent see her particular identities, even those most central to her sense of self, as contingent and particular. She must see that her identity as a rational agent, as a being who needs reasons, is the most important thing about her, and she must then endorse or reject her other identities from this standpoint. Those identities get part of their normativity from this rational endorsement. Furthermore, the agent needs to engage in this process periodically to make sure she can still endorse those identities. So those particular identities will always be contingent and “on probation” for her.

Thus, the man in Williams’ example need not have one thought too many at the moment of rescue. But if his action is to be justified, it must have sprung from an identity that itself was at some point questioned and endorsed from the standpoint of reason. The man at least needs to have one thought too many about his whole husband-identity periodically.

On the Integrated Standpoint View, however, we don’t need to see the moral agent as distancing herself from all her identities in general. We can show why the agent must value humanity from within her particular identities. The distancing that is already built into a particular identity is all that’s needed. Thus, the agent need not see her identities as contingent, partial, or lacking in necessity. A good husband, for example, may never question the necessity of that identity itself, or think that it requires a more universal rational justification. He doesn’t need to see that as a human he needs some identity or other, and this one is pretty good, so he’ll stick with it. As a good husband he will value his wife as a person, respect her humanity, and

distance himself from his own desires to a certain extent: staying up to help his wife when she's sick even though he's tired, sticking around to complete an argument even when he wants to escape, jumping into the water to save his wife even though he's scared. All of that already is a valuing of the capacity for choice, a valuing of humanity. Sometimes a more universal consideration will create a conflict: suppose the husband could reach his drowning wife more quickly by knocking someone else into the water. We could say the husband-identity requires one thing, but the moral identity requires the opposite. But is this really an accurate description? A good husband values his wife as herself *and* as a person. His recognition of the value of others, and his desire not to hurt others, is *part of* his being a good husband. *Qua* husband he wants to save his wife *and* he doesn't want to hurt anyone else. It's not that if he weren't a husband he wouldn't see any problem with knocking people into the water for whatever reason, but rather that his husband-identity is sufficient to motivate him not to knock people into the water, and he doesn't need to step outside of that identity in order to see the general value of humanity. That's why the Integrated Standpoint View can withstand alienation objections much better than the Separate Standpoint View.

Conclusion

If the Integrated Standpoint View is so much stronger than the Separate Standpoint View, this raises the question of why Korsgaard was drawn to SSV to begin with. The answer, I think, is simple: we tend to assume that in order to show that a principle is truly rational, we must be able to demonstrate its rationality from a perspective that is not itself shaped by anything contingent. We must be able to step away from all of our particular concerns, from the things we

happen to care about, and see from that point of view that this principle is still binding. Why do we think this? Because we assume that the paradigmatic rational principles – the rules of logic and mathematics – are grounded in that way, so moral principles must be grounded the same way if they’re going to be similarly rationalized. But that picture of logical and mathematical principles is itself highly suspect, and Korsgaard herself recognizes that in various places (1996, pp. 235-235; 1997, pp. 248-249). When she is able to resist that picture, the view she develops is strong, bold, and an important contribution to ethical theory.

Acknowledgments

I am very grateful for helpful comments from Nancy Bauer, Sarah Buss, Norman Daniels, Lara Denis, Erin Kelly, David Macarthur, Jon Mahoney, Lionel McPherson, Gary Watson, and audiences at Auburn University and Kansas State University.

References

Cullity, G. and B. Gaut (eds), 1997, *Ethics and Practical Reason*, Oxford University Press, Oxford.

Engstrom, S. and J. Whiting (eds), 1997, *Aristotle, Kant, and the Stoics: Rethinking Happiness and Duty*, Cambridge University Press, Cambridge.

Kant, I., 1997, *Groundwork of the Metaphysics of Morals*, translated by M. Gregor, Cambridge University Press, Cambridge.

Korsgaard, C., 1996, *The Sources of Normativity*, Cambridge University Press, Cambridge.

Korsgaard, C., 1997, “The Normativity of Instrumental Reason” in Cullity and Gaut 1997, pp. 215-254.

McDowell, J., 1997, “Deliberation and Moral Development in Aristotle’s Ethics” in Engstrom and Whiting 1997, pp. 19-35.

McDowell, J., 1998a, *Mind, Value, and Reality*, Harvard University Press, Cambridge, Mass.

McDowell, J., 1998b, “Two Sorts of Naturalism” in McDowell 1998a, pp. 167-197.

McDowell, J., 1998c, “Virtue and Reason” in McDowell 1998a, pp. 50-73.

Williams, B., 1981a, *Moral Luck*, Cambridge University Press, Cambridge.

Williams, B., 1981b, “Persons, Character, and Morality” in Williams 1981a, pp. 1-19.

Wittgenstein, L., 1953, *Philosophical Investigations*, translated by G.E.M. Anscombe, Macmillan Press, New York.