

# NATURAL MOTIVES AND THE MOTIVE OF DUTY HUME AND KANT ON OUR DUTIES TO OTHERS

**CHRISTINE M. KORSGAARD**

Harvard University

Christine\_Korsgaard@harvard.edu

**ABSTRACT.** Hume and Kant disagree about the motives involved in the performance of our duties to others. Hume thinks that natural virtues such as benevolence are best performed from “natural” motives, but that there are no natural motives for the performance of the “artificial” virtues, such as justice and fidelity to promises, which are performed from a sense of duty. Kant thinks all duties should be done from the motive of duty. In this paper, I examine the roots of the disagreement. If by a natural motive Hume means an intention that can be described without using normative concepts, Kant would deny that any adult human motives are “natural,” for all involve the thought that something is a reason. But Hume also seems to imply that being motivated to benevolence and self-interest is “natural” in some way that being motivated to keep our agreements is not. I trace this difference to differences in the two philosophers’ conceptions of action. Hume’s conception of action does not allow for genuinely shared action, while Kant’s does. For Kant, being motivated to keep our agreements is just as natural as being motivated to do others good.

## **1. Introduction**

Hume and Kant notoriously disagree about the motives that are, or should be, involved in the practice of our duties to other people. Hume believes that a certain category of duties, those associated with what he calls the “natural virtues,” are generally practiced from what he calls “natural motives” (T 3.2.1,478; T 3.2.6,531).<sup>1</sup> I am going to question this notion, but I will define it for now, as I think Hume does, as motives that we ordinarily have independently of morality. These motives do not involve moral thoughts, or desires whose content must be specified in terms of moral or normative concepts. So when we are moved by these desires, we are not trying to do our duty, or the right thing, or what is virtuous, or what is owed. Many of the natural virtues that Hume identifies, such as prudence, temperance, greatness of mind, cheerfulness, and so on, redound pri-

marily to the benefit of the person who has them, but one of the most important natural virtues, benevolence, is a ground of our duty to be helpful and kind to others.<sup>2</sup> And Hume asserts that:

Tho' there was no obligation to relieve the miserable, our humanity would lead us to it; and when we omit that duty, the immorality of the omission arises from its being a proof, that we want the natural sentiments of humanity.  
(T 3.2.5,518)

On the other hand, Hume believes that the duties associated with what he calls the “artificial virtues” – most notably those of justice and fidelity to promises – are generally motivated by the sense of duty. Kant accepts two distinctions that each captures some features of Hume’s distinction between artificial and natural virtue – namely the distinction between perfect and imperfect duties, and the distinction between strict and broad obligations. But Kant of course believes that all morally worthy actions are done “from duty” (G 4:397).<sup>3</sup>

In this paper I argue that the ground of this disagreement is different than philosophers have traditionally supposed. On the surface, the disagreement appears to be a matter of substantive moral judgment: Hume admires the sort of person who rushes to the aid of another from motives of sympathy or humanity, while Kant thinks that a person who helps with the thought that it is his duty is the better character. While a moral disagreement of this kind certainly follows from their views, I will argue that the source of the disagreement lies elsewhere, namely in their different conceptions of action and motivation. This difference leads in turn to a surprisingly deep difference in their conceptions of our relation to other people, and of what it means to *interact* with other people. It is his conception of human interaction that leads Hume to think that benevolence is natural while there is something artificial about our motives to act justly and to keep our promises. For Kant, on the other hand, no form of adult human motivation is “natural” in Hume’s sense – all adult human motivation involves the agent’s use of non-natural concepts such as “law” or “reason.” But Kant’s theory of interaction grounds another sense in which it is just as “natural” to be motivated to keep our promises and agreements as it is to be motivated to help each other out when we are in need.

## 2. Natural and artificial virtue

I begin by laying out the main elements of Hume's distinction between natural and artificial virtue, as I understand it. A natural virtue is grounded in our approval, that is, the moral sense's approval, of some naturally occurring motive. Characteristically, a person who exhibits that virtue in action is directly motivated by the natural motive in question. For instance, as we have seen, Hume thinks a benevolent person is motivated by his "humanity" – by which Hume presumably means his desire to help others.<sup>4</sup> A person is motivated to do a naturally virtuous action by a sense of duty only if he lacks the natural motive, and hence the natural virtue itself. As Hume puts it:

But may not the sense of morality or duty produce an action, without any other motive? I answer, It may: ... When any virtuous motive or principle is common in human nature, a person, who feels his heart devoid of that principle, may hate himself upon that account, and may perform the action without the motive, from a certain sense of duty, in order to acquire by practice, that virtuous principle, or at least to disguise to himself, as much as possible, his want of it. (T 3.2.1,479)

Importantly, this means that a person who actually *has* the natural virtue need not be aware of, or especially attentive to, the moral character of her actions. She is not, or need not be, motivated by thoughts about what is obligatory or right. Therefore she does not, or at least does not necessarily, think of the helping action as *owed* to the other.

Hume's discussion of the artificial virtues begins from his observation that there are no natural motives for performing the actions required by these virtues. For example, he tells us, just actions are often contrary to the direct interests of those who perform them, the interests of the public, the interests of third parties, and even the interests of the person to whom they are owed (T 3.2.1,480-482). So we cannot appeal to self-interest or to public or private benevolence to explain what motivates us to perform them. But although just actions may not be useful taken one at a time, the *system* of justice is useful or even essential to the workings of society. Because of this, Hume thinks, we were originally motivated by self-interest to support the establishment of a system of justice. And in a small society, we may also be motivated by self-interest to follow the

rules of justice ourselves, both in order to contribute to the maintenance of the system, and because if we were caught violating the rules of justice, others might not allow us to participate in the system and so to reap its benefits (T 3.2.2,499; T 3.2.4,533). Since the moral sense does not especially approve of self-interest, this is not a particularly virtuous motive (T 3.2.2,492). But as society grows and becomes more complex, Hume tells us, we may “lose sight of that interest, which we have in maintaining order” when tempted by the profits of some immediate injustice. Self-interest will no longer sustain our commitment to justice. But even when some particular unjust act might be beneficial to us, according to Hume, “sympathy with *public interest*” leads us to disapprove of unjust actions in a general way (T 3.2.2,499-500). This leads us to disapprove of *ourselves* when we perform them. And that in turn leads us to avoid performing unjust actions, from the motive of duty. Thus in a larger society, the sense of duty becomes the usual motive for performing just actions. We do just actions with the sense that they are obligatory, motivated, according to Hume, by our “regard to justice, and abhorrence of villainy and knavery” (T 3.2.1,479).<sup>5</sup>

### 3. Categories of duty

Hume’s distinction between natural and artificial virtue maps roughly on to the traditional distinction between imperfect and perfect duties or rights. This distinction came into modern philosophy from the Roman Law tradition by way of the work of Hugo Grotius.<sup>6</sup> Although the distinction has been drawn in different ways, generally speaking it goes something like this: when you have a perfect duty you owe some particular action or omission to some particular other person (perhaps to every other person). When you have an imperfect duty, there is some general *kind* of action you are supposed to perform, but you have leeway about exactly what or how much you do, to whom, and on what occasions. You are criticized if you fail to perform your perfect duties, but not ordinarily praised for doing them. You are, however, praised for fulfilling the imperfect duties, especially if you do so to an extraordinary degree – although views differ about what exactly this means.<sup>7</sup> Traditionally, the duties of justice are placed in the category of perfect duties, and the duties of benevolence and charity are placed in the category of imperfect duties. Duties of justice are owed to particular persons,

who are the possessors of corresponding rights; offerings of charity cannot be claimed as a right by any individual, even though each of us has some obligation to practice charity. Claims of justice are upheld by law, while charity is a private matter. We are praised for charity but not for compliance with laws and contracts – rather, we are criticized if we fail to comply with them.

Kant takes this distinction over from the tradition, but complicates it with an additional distinction, between strict and broad duties.<sup>8</sup> Strict duties are directly required actions; broad duties are duties to have certain maxims or ends. Strict duties are all perfect, and may be legally enforced. The requirement to have a certain maxim or end may find expression in either perfect or imperfect duties, depending on whether the performance of a particular action is absolutely required by the end or merely one of many possible ways of promoting it. In Kant's philosophy, the duties of justice and fidelity to promises are strict and perfect duties, while the duty of beneficence is broad and, except perhaps in certain emergency circumstances, normally imperfect – there are many ways we can help others, among which we may choose.

These distinctions have some intuitive force. Many people think that there is some difference between the duties of justice and charity of roughly the sort I have described. And many people would also associate it with a motivational difference of the sort Hume associates with the natural/artificial distinction: the right or at least best way to perform charitable acts is from motives of benevolence and humanity, while just acts must be performed from the sense that they are owed. A familiar objection to Kant's notorious sympathy example in the first section of the *Groundwork* is that someone who does an act of charity or friendship from a sense of duty is inferior to someone who does it from motives of natural sympathy or humanity (G 4:398-399).<sup>9</sup> On the other hand, someone who pays his debts from humanitarian motives seems to be missing an important point – paying your debts is not just a way of being kind to people.

These familiar ideas posed a problem for the rationalists and sentimentalists of the eighteenth century. According to traditional rationalist views like those of Samuel Clarke and Richard Price, rightness or requiredness is a property of actions, and when we recognize that property we are motivated to do them by that fact, and so by the motive of duty.<sup>10</sup> This makes it seem as if all duties are strict or perfect and the motive for performing them is always the

motive of duty, and that gives rise to a puzzle about the status of charitable or benevolent actions. For if every action either has or lacks the essential property of requiredness, it is not clear how we are to make room for the leeway we supposedly have about the amount of charity we practice and the particular occasions on which we choose to practice it. It seems as if we have only two options: either the action is required and so obligatory, in which case it we have no leeway about whether to do it, or it is not required, in which case there is nothing morally good about performing it.<sup>11</sup>

On the other hand, sentimentalist views like Francis Hutcheson's, according to which justice is the practice of "universal calm benevolence," leave it a mystery why any action should be regarded as obligatory or owed rather than merely as better or worse, kinder or less kind.<sup>12</sup> Some contemporary utilitarians have acknowledged this implication of utilitarian accounts of rightness.<sup>13</sup> Hutcheson, who thought of the motive of duty in much the same way Hume does, even suggests that people who act from duty are self-interested, since they are trying to achieve the pleasure of self-approval.<sup>14</sup> Hutcheson himself ends up recommending that we jettison "ought" as a "confused word," which really refers only to justifying and not to motivating reasons. In other words, to say that someone ought to do something is only to say that we approve of his doing it, not to mention a possible motive.<sup>15</sup> Hume's distinction between the natural and artificial virtues is a brilliant attempt to solve this problem, by making a place for something like strict or perfect duties – duties that are rigorous, owed, and done from the motive of duty – within a sentimentalist framework.

#### **4. Motivation and the artificial virtues**

But I want to raise a question about Hume's account of the artificial virtues. What exactly does Hume mean to deny when he denies that there is a natural motive to, say, repaying loans or keeping promises? In setting up his argument, Hume offers us the following case:

I suppose a person to have lent me a sum of money, on condition that it be restor'd in a few days; and also suppose, that after the expiration of the term agreed on, he demands the sum: I ask, *What reason or motive have I to restore the money?* (T 3.2.1,479)

Hume goes on to say that when human beings are in a “rude and natural state,” the answer cannot be “my regard to justice, and abhorrence of villainy and knavery” (T 3.2.1,479). For, he says, “one in that situation would ask you, *Wherein consists this honesty and justice, which you find in restoring a loan, and abstaining from the property of others?*” (T 3.2.1,480). Now it seems initially plausible to suppose that by “in a rude and natural state” Hume means “prior to the institution of justice and property,” since he proceeds to explain how the institution of property arises. Then his point seems to be that motives like “because it is just” or “because I promised” cannot exist in that state, since they depend on institutions and conventions that do not yet exist. But if that is what he means, then he has set up his example very badly. For if a person has lent me money, then there must already be an institution of property, and a convention of lending, which in turn depends upon a convention of promising, all in place. On the other hand, until the institutions of justice and promising are established, there are *no* actions that can be described as “repaying a loan” or “keeping a promise,” so of course there can be no motives for doing these things. So what is Hume denying, when he denies that there are natural motives for just actions?

Perhaps what Hume has in mind is that we can describe the very same actions that we do once the institutions of justice exist, but in naturalistic terms that do not make any reference to the institutions of justice. And when we think of just actions under these alternative descriptions, then we can see no natural motives for performing them. At one point Hume himself seems to suggest this interpretation, for he suggests that the naturalistic description of one just action – the return of property to its owner – is roughly “to restore objects to those who have had first or long possession of them, or have received them by the consent of those who have had first or long possession of them” (T 3.2.6,528). Hume says that the moral sense does not naturally approve of this, and that this shows that justice is not a natural virtue. So perhaps Hume thinks there is no natural motive to just actions when those actions are thus naturalistically described. But there is a difficulty here, for the description of an action either includes the agent’s intention, or it does not, and there are problems either way.

If the description of the action includes its intention, then one may object that Hume has *not* produced a naturalistic description of *the same* action, since returning the object to the person who first acquired it or his successors, is not, or rather is only incidentally, what the person who returns property to its owner intends to do. As Hume himself is eager to emphasize, the exact conventions that determine property are extremely arbitrary, and seem to be based as much on dim psychological associations as anything else (T 3.2.3,501-513). If the conventions that determine property were different, the just person would not return things to their first possessor or his successors – his intention is to do what justice requires, not to return things to their first possessor. More generally, if the description of an action includes the agent’s intention, then there cannot be a wholly naturalistic description of a just action, for it is essential to a just action that it be done with intentions that cannot be specified without appeal to concepts like “property” “promise” “justice” or “right.” If we assume that what renders a motive “non-natural” is that it cannot be described without reference to these normative concepts, then the motives of just actions *must* be non-natural.<sup>16</sup>

On the other hand, Hume might only mean that the moral sense does not approve of the mere outward act of returning things to their first possessor or his successor, leaving the agent’s intention aside. But that would not give him a way to distinguish between the natural and artificial virtues, since as Hume himself emphasizes, the moral sense *never* approves of mere outward acts (T 3.2.1,477). A wholly external description of a benevolent action, if that would be one that made no reference to the benevolent desires of its performer or the neediness of its object, would offer as little for the moral sense to approve of as a wholly external description of a just action does. For consider a parallel. The conventions of justice happen to single out the first possessor or his successors as the owner of property. Well, suppose that the criterion of neediness happens to single out “George.” If I describe my benevolent action simply as “giving money to George” then I have no natural motive to perform it. His being “George” is certainly no reason to give him money. So what is the difference here?

Now it may be objected that I am being petty in focusing so much on the way Hume has set up his example. Okay, so we can concede that Hume has constructed his example badly. His point (the



objector will say) is just that there is no such thing as a just action, and no naturalistically describable item that corresponds to a just action, prior to the invention of the institutions of justice. But naturally virtuous actions *can* be described naturalistically, simply as enactments of the desire to help, and when we do describe them naturalistically we approve of them without further ado. Hume is only telling us that the motives to justice and fidelity to promises came into being along with the institutions of justice and promises themselves. So says the objector.

But actually Hume *does* mean more than this. To see this, consider that there are two different questions we might be asking when we ask for the motive for, say, keeping a promise. Suppose you see me giving a book to someone, and you ask me why I am doing that. I might say I am doing it because I promised that I would give it to him. Essentially what I have done is take an action described intentionally but naturalistically – giving a book to a certain person – and redescribed it in an obviously normative way – as the keeping of a promise. But if what you ask is why I am going to keep my promise, or why in general I keep my promises, you are asking a different question – a question that, in ordinary circumstances, most of us would find a little peculiar. Most of us believe that “I promised” is a sufficient answer, because it is a decisive and sufficient reason; one hardly needs a further reason for keeping a promise. But Hume does think that there is a kind of further reason, namely, that the moral sense disapproves of those who do not keep their promises, and that we will therefore be guilty of villainy and knavery if we do not keep ours.<sup>17</sup> This shows up in the fact that Hume runs the two questions – “why are you giving that person money?” and “why do you repay loans” – or rather the answers to these two questions, together. For instance at the beginning of the discussion of the artificial virtues, continuing a passage I quoted earlier, he says:

I ask, *What reason or motive have I to restore the money?* It will, perhaps, be said, that my regard to justice, and abhorrence of villainy and knavery, are sufficient reasons for me, if I have the least grain of honesty, or sense of duty and obligation.  
(T 3.2.1,479)

Notice the way Hume throws together, without comment, “my regard for justice” and “my abhorrence of villainy and knavery.”

Actually, there is room for a difference between saying that I am motivated by the thought “it was a loan” – that is, by my regard for justice – and saying that I am motivated by the thought “I will be a knave if I do not return a loan” – that is, by my abhorrence for villainy and knavery. If “my regard for justice” means my regard for the requirements of justice, the thought that “it was a loan” might seem to be enough of a motive, without the further thought that only knaves disregard justice. Hume sounds as if he thinks that these two motives are inseparable, that is, that a regard for justice cannot motivate us directly, without the mediation of an abhorrence of villainy and knavery, an abhorrence that comes from the moral sense.

And of course that *is* what he thinks. You might be tempted to protest that the reason you will be a knave if you do not repay a loan is that a non-knavish person *would* be sufficiently motivated to repay a loan by the fact that it was a loan, or by a desire to repay his loans. The moral sense tells us that people who lack these motives are knaves. But now Hume will reply that this isn’t right, and this is just his point, and that this, after all, is what he means by calling the virtue artificial. My judgment that I will be a knave if I don’t repay loans is *not* based on a judgment that there is a certain natural way of being motivated – say, by the desire to return loans – of which the moral sense approves. The moral sense does disapprove of unjust actions, but not because they show the want of some natural motive – rather, because they are contrary to the public interest once the system is set up. Therefore the only possible motive for performing them springs from that disapproval itself.

That is Hume’s view. The point I am making in response to it here is that there appears to be conceptual space for a kind of motivation that stands between Hume’s own categories of natural and artificial motivation. For Hume’s argument in fact trades on two slightly different senses of non-natural: motives can be non-natural in the sense that they cannot be described without reference to normative notions such as justice and property, or they can be non-natural in the sense that they are dependent on the moral sense. And it is conceptually possible to tease these two senses apart. We could hold that once justice and promises exist, people have desires to conform to them: desires to do what justice requires and to keep our promises. These desires, if they existed, would be non-natural in the sense that they would only be describable in terms of concepts drawn from the conventions of justice and promising, but they might still be

natural in the sense that they could motivate us independently of the moral sense. If there were any such desires, the moral sense would certainly approve them. But in that case we would not always do just actions from the motive of duty, as Hume understands it. Rather, one would be motivated to act justly by the sense of duty only if one lacked the natural desire to conform to justice and keep one's promises, just as one is motivated to act benevolently by the sense of duty only if one lacks the natural desire to help. It is clear that this possibility is not on the table for Hume, so it must be that Hume thinks that *even* once the artifice of justice is established, *there are no such desires*. That is, Hume must think that we *cannot* be motivated by the desire to return loans or keep promises *independently* of thoughts about how knavish we would be if we don't. Whereas Hume thinks we can be motivated to help people without having any thoughts at all about how knavish we will be if we don't – for we might just naturally want to help.

This actually fits what Hume says when he sets up his argument. He imagines that the institution has been established, since, as he describes the case, his protagonist has been given a loan, but has no natural motive for returning it. So Hume has not, after all, set his example up in a clumsy way, but said exactly what he meant.<sup>18</sup> But we have cleared Hume of the charge of setting up his case in a clumsy way at a cost. Hume's reason for ruling out the desire to keep promises and repay loans as impossible "in a rude and natural state" is not, we now see, that they must be described in concepts drawn from institutions and conventions that do not exist in a rude and natural state. Rather, Hume thinks that we cannot have desires to keep promises or return loans (independently of thoughts about our knavishness if we lack them) *even* once the institutions exist. But now it is no longer clear that Hume has given us any reason to believe that considerations like "I promised" or "It was only a loan" cannot motivate us directly. Or rather, the only reason he has given us is that the desire to return a loan and the desire to keep a promise *are not forms of benevolence or self-interest*.

But couldn't there be natural motives other than benevolence or self-interest for following the rules of justice? Now you may object that we cannot create natural motives to follow rules simply by making those rules. Our original motive for *making* the rules of justice was self-interest, but as Hume has pointed out, self-interest cannot explain why we *follow* the rules when confronted

with the temptations of free-ridership. So we need some explanation of how people are motivated to follow the rules even in these circumstances, and that is why Hume moves to the explanation in terms of the moral sense: he thinks we are motivated by the prospect of self-disapproval. But there is an important assumption behind this sort of argument: namely, that people are not naturally motivated to keep *agreements*. For once the institutions of justice and promising exist, promises, loans, and transfers of property, are in fact just *forms* of agreement. This is especially obvious in the case of a promise, which may be regarded simply as a rather solemn form of agreement. And there are some philosophers, Hume himself arguably among them, who think the establishment of the institution of property itself is a kind of agreement.<sup>19</sup> Perhaps it is not surprising that philosophers who assume that all motivation is self-interested cannot imagine how we could be motivated to keep an agreement, but Hume does not assume that, since he assumes that we are capable of natural benevolence. So I can put my question this way. Why couldn't we be *naturally* motivated by a desire to keep our agreements with other people in just the same way as we can be *naturally* motivated by the desire to do them some good?<sup>20</sup> According to Hume, people are naturally related to each other only by benevolence or its absence. What conception of human relationships, we might ask, is behind this conviction?

## 5. Kant on the nature of action and the motive of duty

I have been raising the possibility that once the institutions of justice and promising exist, we might have natural motives for keeping promises and repaying loans and the like. We might be naturally motivated to keep our agreements. This suggests that I think Hume's artificial virtues could be constructed in a way that makes them look more like his natural virtues. This may seem like a surprising line of argument coming from a Kantian, who thinks all morally good actions are done from duty. You would expect me to argue that Hume's natural virtues should be more like his artificial ones, rather than the reverse. And in one sense you'd be right. To make things clearer, I will now suggest a way of thinking about the comparison between Hume and Kant, one that becomes available to us once we separate the two versions of the motive to artificial virtue that I have distinguished: namely, "because it was a loan" and

“because I will be a knave if I do not return a loan.” In order to explain what I have in mind, I must explain how I understand Kant’s view of what is rather misleadingly called “the motive of duty,” and in particular how I think it arises directly from his concepts of motivation and action. So now we must take a brief excursus into the philosophy of action.

Kant thinks that the motivation to any action involves two factors, an incentive and a principle or maxim. (I’m using “incentive” here for the variously translated “triebfeder.”<sup>21</sup>) As I understand it, on Kant’s view an incentive is a motivationally-loaded or evaluative representation of an object. I am using the term “object” broadly here to include not only substances but also states of affairs and activities. The object may be something perceived, or just conceived as an attractive possibility. You are subject to an incentive when you are aware of the features of some object that make the object attractive or appealing to you. Perhaps the object satisfies one of your needs; or perhaps because of human nature or your own particular nature the object is one you are capable of enjoying. It interests you, it arouses the exercise of your faculties, it excites your natural curiosity, or it provides some sort of emotional comfort or satisfaction. It doesn’t matter what – something about you makes you conceive this object as appealing or welcome in a particular way. Incentives can also be negative. You may represent an object to yourself as painful or threatening or disgusting, or in some other way unwelcome.<sup>22</sup> The incentive is what presents a certain possible action to the mind as eligible. We can say loosely that desires and inclinations are incentives, and I will do so here, to facilitate the comparison with Hume, but the more correct thing to say is this: to attribute an occurrent desire to someone is to record the fact that he is under the influence of an incentive.<sup>23</sup>

Now here is where the main difference between Hume’s and Kant’s accounts of action comes in. Hume thinks that actions are essentially movements that are caused by our desires, or, in Kant’s language, incentives.<sup>24</sup> To be motivated is essentially to be caused to move by a desire (in light of your beliefs). Kant thinks that desires and other incentives operate on us causally but do not directly cause our movements. If they did, the movements would not be actions. A desire for food, after all, can cause you to salivate. If it also could cause you to go to the refrigerator, then salivating and going to the refrigerator would equally be actions. Action, according to Kant, is

the determination of our own causality, so if we are to count a movement as an action, the movement must be determined by the agent herself, not merely caused by her desires. In other words, an agent must act *on* the incentive, must take it up as a reason for action, by adopting a maxim or subjective principle of acting on it. So every action must involve both an incentive and a principle: that is, something is presented to the agent's consciousness, *on* which the agent then acts. To the extent that the person *determines himself* to intentional movement, he *takes* his desire to provide him with a *reason* for his action; and that is not the same as its directly causing his action. To be *motivated* is to be moved by the thought that some incentive provides you with a reason to act.<sup>25,26</sup>

Since action is the determination of our own causality, on Kant's account, to deliberate is to ask whether a maxim formulated in response to some incentive can serve as law by which I may determine my own causality. One criterion for such a law is that the act I propose to do will in fact enable me to cause the end I propose to bring about. This gives us what Kant calls the hypothetical imperative or principle of instrumental reason. But is not enough for Kant's concept of action to ask whether my maxim can serve as a hypothetical imperative, a law correctly connecting a means to an end. A hypothetical imperative can determine my movements only if I am inevitably determined to move by the end. And if I have no choice about whether to pursue the end, if I am bound to the end by the bare fact that I have a desire for it, then my movements are not self-determined – they are ultimately determined by the end. This is indeed how Hume conceives of acting for an instrumental reason. For Hume, the belief that a certain action will promote my end serves as a kind of conduit by which motive force travels from my desire for the end to the thought of taking means, creating a desire to take the means.<sup>27</sup> But Kant thinks that I am only self-determining if I also determine myself to adopt the end. In other words, I am only self-determining if I adopt the entire maxim as a law determining my causality. So I must ask whether the entire maxim is fit to serve as a law, that is, whether pursuing this sort of ends by these sorts of means can serve as a law. In other words I must ask whether my maxim is in accordance with the categorical imperative, which says to adopt those maxims that have the form of laws.<sup>28</sup>

To avoid confusion, it is important to keep in mind that there are two ways in which a maxim can be said to qualify to be a

universal law. In some cases, when we test a maxim by the categorical imperative, we find that it *can* serve as a law. In that case, the action is permissible. Notice that in this case the incentive we act upon is natural, although the principle under which we act is still ultimately the categorical imperative. In other cases, when we test a maxim, we find that it cannot serve as a law. In this kind of case, we conclude that the maxim of doing the opposite action *must* be a law, and now we have a duty. For instance, if you are tempted to break a promise, test the maxim, and find it rejected, then it is your duty to keep your promise.

Kant thinks that in this kind of case, we are motivated to keep the promise. But if a motive always involves both an incentive and a principle, what in this case is the incentive? That is, what presents the action to the mind as an eligible thing to do? Kant's answer is that the incentive is what he calls "respect for law." To be motivated by respect for law is to be motivated to do something directly by the thought that the maxim of doing it has the force of law, in the sense that it must be a law. In other words, the categorical imperative itself generates an incentive for doing what it says we must do. When we act on this incentive we are said to act "from duty." And when Kant tells us that pure reason can be practical, this is the fact he is referring to – the fact that the principle of pure reason can generate an incentive of its own.

Of course Kant does not mean to suggest that we act from moral incentives only in cases where we are tempted to do the wrong thing. There is no reason to doubt that as we come to know what sort of things the moral law requires of us, we will develop a set of standing moral incentives. And this will include considerations like "I promised" and "it was a loan." The motivational force of such considerations depends on the categorical imperative, in the sense that when we act on them we are moved by respect for law.

## **6. Kant and Hume on the motive of duty**

Now for the comparison. In one way, we can now see that Kant's account closely parallels Hume's account of the operation of the motive of duty, at least in the case of the artificial virtues. In Hume's account, considerations like "I promised" and "it was a loan," or the corresponding desires, get their motivational force from the fact that the moral sense declares that we are knaves if we are not

moved by them. In a similar way, in Kant's account, considerations like "I promised" and "it was a loan" get their motivational force from the fact that the categorical imperative declares such considerations to have the force of law. In the case of these kinds of considerations, according to both philosophers, *motivational force is derived from normative force*, rather than the reverse. That's what the two views have in common. It is in this sense, ultimately, that the motivation involved in acting from duty is not natural – it is motivation that is normative in origin, motivation produced by normative thoughts.

In another way, though, the two accounts are different, because the relation between the general moral principle or faculty and the particular considerations that get their motive force from it is different. In Kant's account, the categorical imperative tells us that certain maxims have the force of law, and have it by virtue of their form. This is supposed to show us that the reasons expressed in those maxims are to be treated as final and sufficient reasons because they have that form. Respect for law provides a motive for doing them, because they *are* laws. So it is not that you have an extrinsic purpose – that of doing your duty – and the categorical imperative tells you that performing these actions is a *means* to doing your duty. It tells you that performing certain actions and adopting certain ends *is* your duty. Because the maxims of these actions have the form of law, respect for law is an intrinsic motive for doing them.<sup>29</sup> It is important to emphasize this, because the usual criticisms of acting from duty, especially in the case of helping people, are based on the idea that your real purpose must *either* be to help the other person or to do your duty. But in fact, what the categorical imperative tells you is that your duty *is* to make helping the other person *your purpose*. Similarly, when you keep a promise, it is not because you have some further purpose – doing your duty – that keeping your promises serves. Your duty is to keep your promise for its own sake. The categorical imperative tells you that a consideration like "I promised" is *itself* to be treated as a reason with the force of law, an unconditional reason.

And it is not clear that Hume can agree with this, or anyway that he does. Consider again his description of the operation of the moral motive in the case of the *natural* virtues:

When any virtuous motive or principle is common in human nature, a person, who feels his heart devoid of that principle,



may hate himself upon that account, and may perform the action without the motive, from a certain sense of duty, in order to acquire by practice, that virtuous principle, or at least to disguise to himself, as much as possible, his want of it. (T 3.2.1,479)

Here it seems clear that Hume is *not* thinking that the moral sense transforms the consideration that someone needs help into a *final and sufficient* reason for action. It is more that your desire to avoid self-hatred or self-disapproval gives you an *instrumental reason* for helping others. You help people in order to avoid self-hatred or self-disapproval. As Hume's predecessor, Hutcheson, had pointed out, it appears as if acting on the motive of duty so conceived is a form of self-interested motivation.

On Hume's theory, do we act justly only to avoid our own disapproval? The question seems to leave Hume with a dilemma. On the one hand, Hume might just accept this implication: even the just person never keeps a promise or returns a loan for its own sake, but as a means of avoiding self-disapproval. On the other hand, he may deny that, and say that the moral sense tells us that considerations like "I promised" and "it was a loan" are themselves, all by themselves, to be treated as reasons for action, and not just as a means of avoiding self-hatred. But then it is not clear why the moral sense shouldn't also tell us that "he needs help" is to be treated as a reason for action, *irrespective of whether I have a desire to help him*, and not just as a means of avoiding self-hatred. If Hume chooses the second alternative, there would be little difference between Hume and Kant, since he would in effect be adopting the view that the moral sense is not merely a faculty of approval and disapproval, but a practical faculty capable of giving rise to reasons through its own legislation. That is, he would have transformed the Humean moral sense into the Kantian will.<sup>30</sup>

But now I want to approach the question from the other side. So let's ask more directly why Kant draws a different conclusion about the duties of beneficence than Hume does. Why is it better if we do them from moral motives? We have seen that Hume thinks an artificial motive is one whose motivational force depends on its normative force. Once that is in place, it is clear that part of the answer is that Kant thinks that, in adult human beings, motivational force is *always* derived from normative force, in the following sense: incentives do not motivate us, unless we decide they have the status of *reasons*, that is, unless we adopt the maxims of acting on them as laws. But that is only part of the answer, because what I have just

said applies to permissible actions as much as to required ones. When we decide that a maxim *can* serve as a law – not that it must, but that it can – we are deciding that a certain natural incentive can count as a *reason* to perform a certain action. So although in Kant’s theory there is a sense in which motivational force always follows from normative judgment, it does not follow that every action is done from the *incentive* of duty: permissible actions are still done from natural *incentives*. And there is a further reason why Kant denies moral worth to benevolent actions done from natural incentives.

Earlier I said that we can distinguish two senses in which you might ask me why I am keeping a promise – that is, doing something that is in fact keeping a promise. You might ask me why I am giving someone a book, and I could say “because I promised,” thus offering a normative redescription of the action. Or you might ask me why I keep my promises – in effect asking me why I treat “I promised” as a reason. We have now seen that Kant’s answer to this second question – why I keep my promises – is that the moral law tells us that promises must be regarded as laws. In a similar way in the case of a beneficent action we might ask, “why are you giving that man a life raft?” to which the answer is “in order to help him” or we might ask “why do you help people?” Now if Hume is proposing that the *virtuous* person’s answer to this *second* question is “because I want to” then we have reached the exact moment where the argument of the first section of Kant’s *Groundwork* begins. For while Kant agrees with Hume that people who want to help others for its own sake are amiable and admirable, he also believes that someone who adopts a maxim of helping others *merely* because he wants to help them or likes to help them does not yet display moral worth.<sup>31</sup> As I have argued elsewhere, Kant’s criticism of the naturally sympathetic person is not that he wants to help others only because it pleases him to do so. The trouble is that he *chooses* to help others *only* because he *wants* to. And this means that he has not yet grasped the deepest reason why we should help people. Since the categorical imperative dictates that we must help each other, people have a claim on our help that is independent of our desires. That reason – that others have a claim on your help – is no more naturalistically describable than thoughts about justice or promising. And beside that reason, any natural desire you have to help people becomes irrelevant, in just the same way that your desire to enrich someone would be irrelevant if you owed him the money.<sup>32</sup>

## 7. Interaction and motives of agreement

I left off my discussion of Hume with the question why Hume couldn't grant that, once the institutions of justice and promising exist, we can have motives for respecting property rights and keeping promises that are natural in the sense that they do not depend on the moral sense. In particular, since promises, loans, and many property transactions, are simply forms of agreement, why can't we be motivated directly by the fact that we have made an agreement, without drawing motivation from the further thought that those who don't respect property rights and keep promises – or perhaps, who don't keep agreements – are knaves? I hope it is now clear that in one sense I think Hume is right about this question and in another he is wrong. I think that Hume is right to deny that our motives for keeping promises and repaying loans are natural and exist independently of morality. According to Kant's account of motivation, *no* form of motivation (or at least adult human motivation) is natural in Hume's sense, for motivation always involves drawing a normative conclusion: at least that your maxim can be a law.<sup>33</sup> And the motivation for justice and promise-keeping is non-natural in a further sense, for in these cases the incentive is derived from the categorical imperative itself: that is, it is respect for law. But in Kant's account the derivation of motivation from the categorical imperative does not take the form of supplying a *further* reason (e.g. that you will hate yourself, or that it is your purpose to do your duty). Rather, the categorical imperative tells us that considerations like "I promised" and "it's his" are *themselves* to be treated as final and sufficient reasons.

But I believe this does not bring out the full force of the difference between Hume and Kant's views. For Hume's view that it is natural to be motivated by benevolence but not by agreement rests not only on a faulty conception of action, but on the faulty conception of interaction that results. In order to explain this point, I turn to Kant's account of interaction.

As I have tried to explain, Kant does not conceive of action merely as a movement caused by something in the agent – say a desire – but rather as a movement determined by the adoption of a principle by the agent. And accordingly, he does not conceive of *interaction* merely as the production of effects by two agents on each other, not even as the mutually deliberate production of such effects.

Instead, he conceives of interaction as shared or joint action – as a set of movements that is determined by a jointly adopted principle. On Kant’s account, people who interact must act together in a literal way – by deliberating together to arrive at a shared maxim that will then govern their conduct towards each other.

This shows up most clearly in Kant’s treatment of exactly the cases we have been talking about – the cases of promising and property – in the *Metaphysical Principles of Justice*. Kant’s theory assimilates these two cases, for Kant conceives all rights on the model of property rights, and that means that he supposes that when we make someone a promise what we are doing is in effect giving that person a kind of property right in an act of our own. To make a promise is to transfer your right over your choice whether to perform the action in question to the other person. Your action, or at least your choice whether to perform it, now belongs to the other and the other can claim it. So if I promise you that I will meet you for lunch tomorrow, a certain act of my will – the decision whether to meet you for lunch – now belongs to you and not to me. That is why it is you and only you who can release me from my promise – because the choice in fact belongs to you.

But there is a problem, Kant claims, about how this can happen. Suppose I offer you a promise. Until you accept my promise, I can always take it back – I am not committed until you have accepted. But if in making my offer I have not yet committed myself, then I have not yet promised, and you have nothing to accept. If a promise consisted of two separate acts, this problem would be insoluble, for there is necessarily a temporal gap between my offer and your acceptance, and however small the gap is it prevents the transaction from happening. The same problem exists and is even more vivid when what is happening between us is a transfer of an ordinary piece of property rather than of an act of my will. How do I transfer my property to you? It cannot be that I first abandon my property and then you pick it up, for if I abandoned it then during the interval anyone could legitimately pick it up – it would be unowned property, which anyone may claim. Yet until I have abandoned it, it is mine, so how can it possibly become yours? If you take it when it’s mine you’ll be stealing it. The gap between the two acts again prevents the transfer. These problems show that promises and transfers cannot be understood as the result of two separate and successive acts. Instead they must involve the formation of a single

common will, that is, the adoption of a shared principle or maxim. You agree to take what I agree to relinquish, I agree to relinquish what you agree to take, and so we form a unified will that the thing in question should be yours and not mine: it all has to happen at once, as a single action, if it is to happen at all.<sup>34</sup> The free choice of this maxim is an act that constitutes our unified will and makes shared action, and so interaction, possible.<sup>35</sup>

Now in one way this account makes it look as if our motives for keeping our agreements are natural. At least, it looks as if, once we regard interaction in the way that Kant suggests, our reasons for keeping our agreements are no more mysterious (or perhaps I should say no less mysterious) than our reasons for carrying out any of our own choices, whether we make them alone or along with others. If we regard interaction as shared action, we don't need a *special* reason for keeping our agreements, any more than we need a special reason for doing the things we've decided to do. But why *should* we regard interaction in this way?

The answer – or at least the part of the answer I am going to discuss in this paper – is that it is required by the conception of others as ends in themselves. To regard another as an end in himself is to regard him as a source of reasons, and this conception is expressed in the fact that, when you deliberate about matters by which you will be jointly affected, you take his reasons as well as your own into account. By “his reasons,” here, I do not mean the things *you* take to be good for him, but rather the things that *he* takes to be reasons. And to explain what I mean by “take into account,” I will make a comparison. Your own incentives appear in your private deliberations as candidates for reasons, and you deliberate by asking whether acting in a certain way on a certain incentive could serve as a law for you. When you and I engage in a shared deliberation, the reasons endorsed by each of us play the role of incentives in the shared deliberation; we ask together whether the maxim of acting on these reasons could serve as a law for *us*. When you act on a maxim that the two of us cannot possibly share – as you certainly do when you unilaterally set aside a principle that we have adopted together – you disregard my reasons, and so treat me as if I were not an end in itself, but merely a tool, or perhaps an obstacle, to ends of your own. So the conception of human beings as ends in themselves requires us to view interactions as shared actions.

The duty of beneficence springs from the same conception of others. Another person is a standing source of reasons, reasons that you cannot just ignore. On a given occasion, your own reasons may outweigh the other person's, which is part of the reason why the duty is normally imperfect. But the reason you help him is not just that you want to: it is that his reasons constitute claims on you in the same way as your own.<sup>36</sup>

## 8. Conclusion

I conclude that there is a sense in which it is just as natural to be motivated by agreement as it is by benevolence. That sense is given by the fact that respect for the law, and respect for humanity, are for Kant one and the same thing. According to Kant, to be motivated by the motive of duty, ultimately, is simply to be motivated by the other person himself, by the sheer normative fact of the other person himself. He's *there*, he's a source of reasons, a fellow deliberator and a standing source of claims. And in one way, Hume agrees with this. For Hume's moral sense is not just the voice of our private approvals and disapprovals. Because of sympathy, and its universalizing effects, it is the voice of humanity itself, internalized in each of us. And when it regulates our conduct, we are in a sense *there* for each other, present to each other, making normative claims. But although Hume's moral theory allows us in this way to be *influenced* by the claims of others, his theory of action leaves us fundamentally divided from one another.<sup>37</sup> For a Humean agent is just a cause, like any other. On Hume's conception of action, we cannot literally act together, we can only have effects on one another. It is not surprising that someone with this conception of action finds benevolence more intelligible than agreement as a source of motivation.

Kant's conception of action allows us to deliberate and act together in a deeper way. The other person literally occupies the space of reasons with us, engages with us in deliberation that leads to actions that we carry out together. So when you do something just because of the other person, just because he is there, because he is a law to you just as you are to yourself, is your motivation natural or moral? The answer is simple: it is moral – you are acting from the motive of duty – but it is also the most natural thing in the world.<sup>38</sup>

## NOTES AND REFERENCES

1. References to Hume's *Treatise of Human Nature* (2nd edition ed. L. A. Selby-Bigge, revised by P. H. Nidditch. Oxford: Clarendon Press, 1978), will be inserted into the text, as follows:

T 3.2.1.478 = *Treatise of Human Nature*, Book 3, Part 2, Section 1, p. 478.

2. That is to say, the fact that benevolence is a virtue is the ground of our duty to be helpful and kind.

3. References to Kant's works will be inserted into the text in the conventional fashion, using the volume and page numbers of *Kants gesammelte Schriften* (published by the Preussische Akademie der Wissenschaften, Berlin), which are found in the margins of most translations. G indicates the *Groundwork of the Metaphysics of Morals* and MM indicates *The Metaphysics of Morals*.

4. There is room for disagreement about how we should understand the relations between motives and desires. Some philosophers would say that to be motivated by what Hume calls "humanity" is to be moved by the *fact* – or by the belief or thought – that another person is in need of help. But since Hume is inclined to think of motives as desires, perhaps we should say that according to Hume "humanity" is a desire to help those in need. I will not be examining this dispute directly in this paper, but it is closely related to the disagreement I do examine here. For a discussion of the disagreement about whether motives are desires or the facts to which those desires are a response, and the relation of that question to the motive of duty, see my "Acting for a Reason," in Korsgaard, *The Constitution of Agency* (Oxford: Oxford University Press, 2008), pp. 207–229.

5. In the second *Enquiry*, Hume does not appeal to this natural/artificial distinction, but some of the same work is done by an idea that Rawls later called "the circumstances of justice" (in *A Theory of Justice*, Cambridge: Harvard University Press, 1971, p. 126). Hume establishes that justice is valued for its utility by arguing that in circumstances in which justice is not useful, it is also not required. If people were infinitely benevolent or goods infinitely available, justice would not be necessary for anyone. If other people were completely malevolent or goods so scarce that not everyone could hope to survive, justice would not be useful to everyone. But under these conditions, Hume think, we are not bound by justice. The conclusion – that the value of justice holds only when everyone can benefit from the adoption of a certain system of conduct, is similar to that of the *Treatise*.

6. In *The Law of War and Peace*, Book I, Section 4. Grotius wrote of perfect and imperfect rights. A "perfect" right is complete – it is enforceable, and it is owed either by everyone or someone in particular to the person who has it. An imperfect right is not complete because it attaches only to the person who has it; there is no clear corresponding duty. At some later point, the distinction shifted so that it was in the first instance between two kinds of duties. I am drawing on the selections from Grotius and the discussion by Jerome Schne-

wind in *Moral Philosophy from Montaigne to Kant*, volume I (New York: Cambridge University Press, 1990). See especially p. 97.

7. The traditional doctrine of supererogation holds that performing certain actions, or perhaps certain numbers of actions, counts as going beyond what is required, though the actions are still morally good. Kant thinks we are praised in general for the performance of imperfect duties, but nothing in particular counts as doing “more than we are required to.” Richard Price sometimes suggests that a particular amount of benevolent action *is* required of us, but that we cannot know exactly what that amount is. Under these circumstances we admire people who do many benevolent actions for erring in the right direction, since the motive of doing so is good. (See pp. 169-172 in the selections from Price in *British Moralists*, Volume II, edited by D. D. Raphael. Indianapolis: Hackett Publishing, 1991.)

8. The categorization of duties into strict and broad is mentioned in the *Groundwork* but not explained until the *Metaphysics of Morals*. In the *Groundwork* Kant at one point appears to equate it with the perfect/imperfect distinction (G 4:424), but in the *Metaphysics of Morals* it becomes clear that this would not be correct, since there are broad perfect duties. All strict duties are perfect, although not all perfect duties are strict; only strict duties may be legally enforced. In the *Groundwork*, Kant says that a perfect duty is “one that admits no exception in favor of inclination,” suggesting that an imperfect duty does admit of such exceptions. In the same place he asserts that his view differs from that of “the schools” because he admits internal perfect duties (G 4:421n.) In the *Metaphysics of Morals*, he classifies certain duties of virtue owed to the self as “perfect,” and it seems natural to classify some of our inner duties of virtue to others – those of respect as opposed to love – as “perfect” as well. The relationship between the two distinctions is somewhat confusing and there is not much general agreement among commentators about how to understand it. It will not be necessary for purposes of this paper to go into it in any detail.

9. I explain what is wrong with the objection in Section 6 below.

10. Price characterizes right and wrong as “real characters of actions...which denote what actions *are*.” I am quoting from the selections from Price’s *Review of the Principal Questions and Difficulties in Morals* in D. D. Raphael’s *British Moralists* (Indianapolis: Hackett Publishing Company, 1991), volume II, p. 133. For Price’s assertion that we are motivated by the sense of obligation, see p. 194.

11. Price, who noticed this problem, suggests that it is the *kind* of action rather than the particular actions that are required of us in the case of the imperfect duties, but it is not really clear why the property of rightness should sometimes attach to particular actions, and sometimes to kinds. (See Price’s discussion in *British Moralists*, volume II, at pp. 169-173). On the other hand, it is arguable that there is no reason why the property of rightness, conceived in this realist manner, should attach to any one thing rather than any other, and therefore that there is also no reason why not. For a discussion of this claim about how intuitionist theories work, see my “Normativity, Necessity, and the Synthetic a



priori: A Response to Derek Parfit” available on my web page at <http://www.people.fas.harvard.edu/~korsgaard/Korsgaard.on.Parft.pdf>.

12. See the selections from Hutcheson’s *Inquiry concerning the Original of Our Ideas of Virtue or Moral Good* in *British Moralists*, volume I, pp. 282–283.

13. For one example see Frances Howard-Snyder and Alastair Norcross, “A Consequentialist Case for Rejecting the Right.” *Journal of Philosophical Research* 18 (1993): 109–125.

14. This is the main ground on which Hutcheson criticizes Shaftesbury. See the selections from *An Inquiry concerning the Original of our Ideas of Virtue or Moral Good*, in D. D. Raphael’s *British Moralists*, Volume I, see p. 305.

15. *Illustrations on the Moral Sense* (Cambridge: Harvard University Press, 1971), p. 140. For the discussion of justifying vs. motivating or “exciting” reasons see pp. 121ff.

16. This of course is not to say that the motive is non-natural in some spooky sense, that it is itself a non-natural entity. From a third-person perspective, we can avoid that implication by a kind of device: we can say that a just action is one motivated by the fact that the agent *thinks* it is just or *desires* to do what is just. There is nothing non-natural about the existence of thoughts and desires couched in non-natural concepts, provided we can give a naturalistic explanation of how the agent came to be using those concepts. And Hume of course thinks that we can do that – that is what his argument is intended to do, to explain how we came to use these non-natural concepts.

17. You might think Kant also supposes that there is a further reason, namely that the categorical imperative says to keep your promises. But this would not be quite right. What the categorical imperative says is precisely that the fact that you have promised *is* a reason (not that there is a further reason), as I will explain below.

18. Although on the reading I propose here it is still not clear what Hume could possibly have meant by “in a rude and natural state.” I am inclined to think that what happened is that because Hume did not think explicitly about the difference between motives like “it was a loan” and “I will be a knave if I don’t repay loans,” he himself confused the idea that no natural motives to justice exist because such motives cannot be described without reference to the institution of justice with very different idea that we cannot be naturally motivated by thoughts like “it was a loan.” In effect, he assumed his own conclusion: that the motivation to repay loans depends on the motivation to avoid knavery. However, the explanation may be simpler. On Hume’s theory, there is presumably a gap between the time when we set up the institution of justice and the time when we begin to think about justice as a moral issue, which perhaps doesn’t happen until society becomes larger. Perhaps the “rude and natural state” occupies that gap in time.

19. Hume describes property as “a convention enter’d into by all the members of society to bestow stability on the possession of those external goods,

and leave everyone in the enjoyment of what he may acquire by his fortune and industry.” (T 3.2.2,489)

20. In the famous passage about the two men rowing the boat together, Hume suggests that people who act on informal agreements are motivated by “a sense of common interest” (T 3.2.2, 490). Hume thinks this is how members of society agree to set up the institutions of property and promising. This is obviously not what I mean when I talk about being motivated by the fact that you have made an agreement. I believe that at least until Butler clarified the relationship between particular passions and self-interest in his Sermon XI of his *Fifteen Sermons Preached at the Rolls Chapel*, Hutcheson and Hume had some tendency to think of all motives as *forms* of benevolence and self-interest. It is tempting to think that this confusion is at work here.

21. The troublesome word *Triebfeder* is rendered differently by Kant’s various translators. It is translated as “incentive” by Mary Gregor, James Ellington, and in the earlier version of Lewis White Beck’s translations of Kant; as “impulsion” by H. J. Paton; as “spring” by Abbot, and as “drive” in Beck’s later translations.

22. For further discussion see Korsgaard, *Self-Constitution: Agency, Identity, and Integrity* (Oxford: Oxford University Press, 2009), section 5.6, pp. 104–108.

23. See Korsgaard, *Self-Constitution* (cited above), section 6.2.5, pp. 120–121.

24. The clearest evidence for this is of course from the sections on free will (T 2.3.1-2), where Hume asserts that there is no difference between human causality and other kind, and, even more importantly for my purposes here, equates a person being responsible for an action with its being caused by something in him. (T 2.3.2,411).

25. For this conception of motivation, see my “Acting for a Reason,” cited above. What I say here – and indeed what Kant says – makes it sound as if only human beings can be motivated and so can be agents, and in my view that is clearly wrong. I address this problem in *Self-Constitution* (cited above), chapter 5.

26. I am calling this Kant’s theory of action, but someone may object that this is only Kant’s theory of autonomous action. Kant himself was initially confused about this, for in the *Groundwork*, at least, he tried to make a distinction between autonomous and heteronomous action, and suggested that a movement caused by a desire would be a heteronomous action. But in fact, according to his own theory, a movement caused by a desire would not be an action at all, for it would be attributable to the desire and not to the person, to his will. Action must be autonomous by its nature. Later, he switched to the view that an action always involves a principle, but some actions are chosen under the wrong principle. I discuss this below in the text. For further discussion of the view that action is essentially autonomous see Korsgaard, *Self-Constitution*, chapters 4–5; for further discussion of the idea of choosing under the wrong principle, see chapter 8.

27. As I have argued elsewhere, this way of thinking of the effect of instrumental reasoning does not get us to an adequate conception of the hypothetical imperative, that is, to a *normative* principle of instrumental reason. See my “The Normativity of Instrumental Reason” in *Ethics and Practical Reason* (ed. Garrett Cullity and Berys Gaut. Oxford: Clarendon Press, 1997), chapter 8, pp. 215–254; and in Korsgaard, *The Constitution of Agency* (cited above), chapter 1, pp. 27–68.

28. There is a question whether and why this is the same as its being fit to serve as a universal law in the sense needed to generate moral conclusions – a law universalizing over all rational beings, treating reasons as having shared normative force for those beings. I have discussed that question elsewhere, and will say something about the grounds of treating reasons as having shared normative force in the discussion of interaction in section 7 below.

29. It’s important here to observe that what Kant identifies as the moral motive is “respect for law” not “respect for *the* law. Translating it the latter way raises the question “which law?” and suggests that the answer is “the categorical imperative.” But that isn’t how it works. Rather, the categorical imperative in effect tells us to respect the lawlike form of any maxim that has such a form, in the sense that by virtue of its form it must be a law.

30. This is essentially what happens in the view of Adam Smith.

31. Actually, it is possible that someone who helps others from natural inclination is acting under the categorical imperative as his principle of choice, but only concluding that it is permissible for him to satisfy his natural desire to help. Such a person would not be following the principle of self-love, but would also be failing to perceive the deeper reason for helping.

32. For a fuller account of this reading of the argument of *Groundwork* I, see my “From Duty and for the Sake of the Noble: Kant and Aristotle on Morally Good Action,” in *Aristotle, Kant, and the Stoics: Rethinking Happiness and Duty* (ed. Stephen Engstrom and Jennifer Whiting. New York: Cambridge University Press, 1996), chapter 7, pp. 203–236; and in Korsgaard, *The Constitution of Agency* (cited above), chapter 6, pp. 174–206.

33. Kant’s account gives rise to difficulties about how we are to understand the motivation and actions of the other animals. For discussion, see Korsgaard, *Self-Constitution* (cited above), chapter 5.

34. These arguments are from *The Metaphysics of Morals*, 6:271–276.

35. For further discussion of this conception of interaction see Korsgaard, *Self-Constitution* (cited above), chapter 9.

36. This is why there is an element of agreement in the Kantian duty of beneficence – it is the duty to promote the happiness of the other *as he sees it*, not just to do good things for him. (*Metaphysics of Morals* 6:388)

37. One of the strangest manifestations of this individualism, in my view, is Hume’s curious conviction that if promising were natural it would be some sort of individual act of the mind (T 516). It is almost as if he thinks that the very fact that promising is essentially relational renders it unnatural.

38. An earlier version of this paper was written as a keynote lecture for the meetings of the Hume Society Meeting in Tokyo in 2004. I am grateful to the Hume Society for inviting me to speak. The paper itself was delivered at Quentin Skinner's Political Philosophy Seminar, at the North American Kantian Society Study Group, as the keynote lecture at the University of Toronto Graduate Student Philosophy Conference, at Ronald Dworkin's Colloquium in Legal, Moral and Political Philosophy at University College London, and at the University of Alabama at Birmingham. I am grateful to the audiences at all these occasions for useful discussion and response.