

- Velleman, David (2000d) 'What Happens When Someone Acts', in Velleman (2000a).
- Wedgwood, Ralph, 'The Aim of Belief', *Noûs-Supplement: Philosophical Perspectives*, 16 (2002), 267–97.
- Williams, Bernard, 'Deciding to Believe', in *Problems of the Self* (Cambridge: Cambridge University Press, 1973), 136–51.
- Zangwill, Nick, 'Direction of Fit and Normative Functionalism', *Philosophical Studies*, 91 (1998), 173–203.

## 10

## Misunderstanding Metaethics: Korsgaard's Rejection of Realism

*Nadeem J. Z. Hussain and Nishi Shah*

### 1. Introduction

Contemporary Kantianism in ethics is often thought of not just as a position within normative ethics but also as an alternative to moral realism. We argue that it is in fact not at all clear how contemporary Kantianism can distinguish itself from moral realism. There are of course many Kantian positions. For reasons of space we have chosen to focus here on the position of one of the most prominent, contemporary Kantians, Christine Korsgaard. Officially our discussion is restricted to her version of Kantianism, though we suspect that the lessons learnt here apply elsewhere.

In our experience, it immediately strikes some as implausible that Korsgaard is actually engaged in metaethics. We grant that there are strains in Korsgaard that suggest an attempt to, so to speak, go "beyond" metaethics. We take up such a reading of Korsgaard elsewhere (Hussain and Shah, 2005*b*). Here we simply accept at face value the way in which she repeatedly introduces the Kantian view as an alternative to realism. Crucially, she emphasizes that the realism of concern to her is, as she puts it, "*substantive* moral realism"—that is, a view with specific metaphysical, epistemological, and semantic commitments. It is "the view that there are answers to moral questions *because* there are moral facts or truths, which

We would like to thank Allan Gibbard, Pamela Hieronymi, Elijah Millgram, Tamar Schapiro, Yonatan Shemmer, Kenneth Stalzer, and Sharon Street for very useful conversations about the issues raised in this paper. We would also like to thank the participants of the Annual Metaethics Workshop, University of Wisconsin-Madison, 2004, for their questions and comments. Thanks to Sarah Buss, Alexander George, Kieran Setiya, David Velleman, Allen Wood, and an anonymous referee for written comments on an earlier draft of this paper.

those questions ask *about*" (Korsgaard, 1996: 35).<sup>1</sup> According to realism, moral requirements must be given "some sort of ontological foundation, by positing the existence of certain normative facts or entities to which moral requirements somehow refer" (Korsgaard, 1997: 218). Not surprisingly she does not want to contrast her own position with uses of the word "realism" that merely mark out a contrast with nihilism—realism merely as the general normative view, that is, that there are correct answers to questions about what we should do (Korsgaard, 1996: 35). For the purposes of this paper we take this as sufficient evidence that (i) she is contrasting Kantianism with the metaethical position of realism and that (ii) she takes Kantianism to be the philosophically favoured position of the two. Our claim is that she fails to show either that Kantianism is different or that it is better than realism.

## 2. The Normative Question(s)

Our general strategy will be to argue that what are supposed to be claims that conflict with realism in fact fail to do so. We will rarely attack the arguments for these claims. What we will attack instead is the argument against realism based on these claims. These claims (and the arguments for them) fail, in general, to undermine realism because Korsgaard fails to show that they actually conflict with realism in the first place. They often fail to conflict because though they may appear to be metaethical claims they in fact are not obviously so and indeed are most charitably interpreted as either claims within normative ethics or normative psychological claims in the philosophy of action, claims compatible with several different metaethical accounts of those same claims including non-reductive normative realism.

We will argue therefore that what explains the failure to distinguish Kantianism from realism is a failure to appreciate all the consequences of the traditional distinction between normative judgements and metaethical interpretations of normative judgements.<sup>2</sup> Thus we begin with a brief review of the differences between normative ethics and metaethics. Within normative ethics, we can distinguish at least two different philosophical tasks. The first is to construct a set of principles that systematize and

<sup>1</sup> Emphases in original.

<sup>2</sup> As we have already noted, one can read Korsgaard as intending to undermine the distinction between normative ethics and metaethics. We take it though that the distinction is supposed to be undermined in part as a consequence of her arguments against realism (and other metaethical views). An argument for the claim that her view is different from, and better than, realism cannot simply presuppose that the distinction has been undermined. For further discussion, see Hussain and Shah (2005b).

ground our correct moral judgements. Utilitarianism and various forms of deontology are examples of such theories, expressing competing conceptions of the fundamental moral principle(s) from which correct judgements of moral rightness and wrongness can be derived.<sup>3</sup>

The second task is to place morality within practical reason, explaining whether we have reason to do what morality demands and, if so, whether these reasons are derived from another branch of practical reason. There are two ways of carrying out this task. One is to argue that it follows from the concept of a reason for action or agency or well-being that an agent always has reason to do what is right and avoid doing what is wrong. The debate about the conceptual possibility of the amoralist (someone who judges an action would be right but sees no reason to do it) is about the success of this strategy of placing morality within practical reason. While many philosophers would label this a debate within metaethics, we place it under the heading 'normative ethics' in order to mark the fact that, whichever side one takes, one will not yet have answered the questions that Korsgaard's stated adversary, the realist, attempts to answer. Recall that Korsgaard's "*substantive moral realism*" is a position with specific metaphysical, epistemological, and semantic commitments. Realism is not a position about the relation between some normative concepts (for example, "rightness") and other normative concepts (for example, "reason for action"), but is a position about the nature of normative concepts in general.

The other way of placing morality within practical reason is to show that moral requirements follow from a substantive conception of practical reasons. How one carries out this strategy depends upon one's conception of practical reason. If one thinks that the aim of practical reason is to maximize an agent's desire-satisfaction, then placing morality within practical reason will entail showing that doing what morality demands maximizes the satisfaction of an agent's desires. But if one has some other conception of practical reason, then showing that morality satisfies desires may be beside the point; instead one might be faced with the task of showing that the demands of morality can be derived from something else, for example, the principles of autonomy. Or, there may be no need to show that morality can be derived from anything at all, if according to one's conception of practical reason, the principles of morality are fundamental principles of practical reason.<sup>4</sup>

<sup>3</sup> There are many options here, e.g. upon investigation, one might conclude that there are no deep, exceptionless moral principles (see Aristotelian theories), and that the best we can do is arrive at more-or-less useful rules of thumb.

<sup>4</sup> Note that in this context the claim that moral principles are fundamental principles of practical reason is a substantive claim about the correct conception of practical reason,

A metaethical account offers an interpretation of the normative claims that are offered as answers to these inquiries (for example, that it is morally right to maximize utility or that one has reason to do those actions that are morally right), aiming to tell us what these claims mean, whether they involve metaphysical commitments, if so, what these commitments are, and whether and how we acquire knowledge of these normative claims. Non-reductive realism and non-cognitivism are examples of positions that give competing answers to these questions. The non-reductive realist usually subscribes to a referential semantics (the judgement that *x* is good expresses a belief that *x* has a normative property), an ontology of non-natural properties (normative properties are non-natural properties), and an intuitionist epistemology (we come to know basic normative truths by non-sensory, rational intuition). Non-cognitivists, on the other hand, usually reject a referential semantics for moral terms. They claim that moral judgements do not express truth-evaluable beliefs in normative facts, but express non-depictive motivational states such as desires, preferences, or emotions. Non-cognitivists are therefore free to accept an ontology restricted to natural properties and to deny that there is an epistemology needed for moral judgements, since moral judgements, being non-depictive and therefore not truth-evaluable, do not aspire to knowledge.<sup>5</sup>

We do count views that argue that there is, in some sense, no way of getting outside of normative thought to explain it, and that therefore no answers to these questions are possible, as doing metaethics. However, this type of quietism, which claims that no metaethical theories are possible, is not equivalent to merely failing to state a metaethical position. One might pursue normative ethical tasks while ignoring metaethical ones, leaving such questions for others to answer. This acceptance of a division of philosophical labour certainly would not commit one to the quietist claim that metaethics is impossible. Quietism is a bold position in need of justification, whereas the decision to pursue normative ethical questions instead of metaethical ones needs no philosophical defence.<sup>6</sup> The point of metaethics is to give an

not an analytic claim about the relation between the concept of a reason for action and the concept of moral rightness. Our thanks to an anonymous referee for *Oxford Studies in Metaethics* for encouraging us to be clearer about the differences between normative ethics and metaethics. For further clarification, see Hussain and Shah (2005b).

<sup>5</sup> We simplify; the non-cognitivist has to give us some account of what we are doing when we claim that we know that murder is wrong.

<sup>6</sup> It is perhaps only relatively recently, *c.* 1903, that philosophers have self-consciously isolated and pursued specifically metaethical questions. Thus, when interpreting ethicists in the history of philosophy who did not explicitly distinguish these questions, we must be very careful not to assume that because of the metaphysical or epistemological sounding labels used to express their views that they are always making metaethical

account of what it is to think a normative thought, or to show that such an account is impossible, not to tell us which normative thoughts to think or to point out which normative thoughts we cannot help but think.

Even in this traditional form of the distinction, normative ethics and metaethics are not completely independent of each other. Since metaethics is an attempt to provide an interpretation of our normative practice, which metaethical theory we end up with will in part be determined by what we think our practices of making normative judgements look like. Furthermore, normative ethics may lead us to think that certain moral claims are true. Ascribing error has its costs and so metaethical theories that allow these judgements to be true will have a defeasible advantage. Similarly certain metaethical theories, reductive realism for example, will entail particular normative claims. One cannot claim that 'right' just means "maximizes utility" without its following that if an action maximizes utility, then it is right.

With this traditional distinction in hand, we turn in the next section to the task of trying to identify Korsgaard's Kantian position by focussing on her insistence on distinguishing her own view from a position she labels 'realism' or 'dogmatic rationalism'. We take the target here to be non-reductive normative realism and ask whether her rejection of non-reductive realism might give us insight into her alternative Kantian view. Our claim is that her central objection to the non-reductive realist reveals the above-mentioned failure to distinguish between different questions about normativity and that the non-reductive realist has a coherent response. This failure to distinguish answers to normative questions from answers to metaethical questions also undermines Korsgaard's attempt to show that her own position is an alternative to non-reductive realism.

We then take a detailed look at her account of instrumental reason. We carefully assess her account of the will, the idea of a constitutive norm, and the role of self-legislation to see if we can identify a positive Kantian position that can respond to the worries about non-reductive realism raised by both her and others. We conclude, however, that it is in fact very hard to see how Kantianism about instrumental reason could represent a position distinct from non-reductive realism.

We finish by assessing whether Korsgaard's constructivism and its account of normative concepts succeeds, as it is apparently supposed to, in establishing an alternative to non-reductive realism. We conclude that, in its

claims. This is why we think it best to avoid simply using the historical labels, such as 'rationalism', 'empiricism', and even 'voluntarism', that Korsgaard uses to describe various ethical theories, as these labels often stand for historical theories in ethics that ran together positions in normative ethics and metaethics.

currently undeveloped state, it does not and that thus in the end Korsgaard leaves us with no distinctive Kantian alternative to non-reductive realism.

### 3. Sources of Normativity

#### 3.1. Sources of Normativity

In this section we argue that in Korsgaard's attempt to delineate "the normative question" in *The Sources of Normativity* (1996) she fails to distinguish the task of placing normative principles and judgements within practical reason from the task of giving a metaethical account of those principles and judgements. This causes her to misunderstand the aims of Prichard and Moore's metaethical views and to reject them on the spurious grounds that they fail to answer questions within normative ethics. We then argue that her own solution to the "normative problem" is infected by this ambiguity, and thus fails to express a distinctive metaethical view, much less one that contrasts with the non-reductive realist views of Moore and Prichard that she rejects.

The failure to distinguish normative from metaethical questions is reflected in a potential ambiguity in Korsgaard's claims to have identified the "source of normativity" or to have "explained normativity". There is a distinction between what *makes* an action wrong or a principle normative, on the one hand, and what *constitutes* the normativity or what the property of being normative itself is, on the other. Thus the fact that brushing my teeth regularly will reduce plaque may *make* brushing my teeth good (for me); however, we do not want to claim, presumably, that the property of goodness itself just is the property of reducing plaque.<sup>7</sup> The ambiguity mentioned above can now easily be seen. There are perfectly understandable senses of these expressions according to which one might well say that one is picking out the source of the normativity of teeth brushing—or explaining why one ought to brush one's teeth—by pointing out that the brushing of teeth reduces plaque. But these claims are best understood as first-order normative judgements about what makes brushing one's teeth good, not as

<sup>7</sup> We are not claiming that the sense of expressions of the form "make *x* wrong" that we are trying to pick out and use here is exhaustive or even central to the ordinary language uses of such phrases. The hope is to use this phrase essentially as a term of art—a now almost standard one—to help keep track of an important philosophical distinction. Note, we do not deny that some such identification of wrongness with what makes things wrong is part of the strategy of certain realists. Our assumption of course is that such an identification is not likely to be part of any metaethical strategy that would be recognizably Kantian. More on this below.

providing a metaethical interpretation of what it means to say that reducing plaque is good, what metaphysical commitments such a judgement involves, or how we come to know that brushing one's teeth is good.

In *Sources*, Korsgaard says that in seeking a philosophical foundation for morality we are not looking for a mere explanation of morality, but for a justification of the claims that morality makes on us (1996: 9; also 16). She says that giving an adequate third-personal explanation of morality, such as might be given by an evolutionary account of morality according to which morally right actions are those that promote the preservation of the species, would not answer "the normative question" because it would fail to justify morality from within the first-personal point of view (14). This suggests that she is seeking to place the claims of morality within practical reason. The normative position that claims that right actions are those that promote evolutionary fitness would be a failed attempt to place the claims of morality since it would not show moral claims to be justified from within practical reason. That some action would promote fitness does not seem at all like a reason to do that action. However, such a position would be an example of a failed theory in normative ethics rather than a metaethical theory.<sup>8</sup>

Korsgaard's description of the "substantive realist" answer to the normative question, however, depicts it as a metaethical position. Then again, the main criticisms that she makes of "substantive realism" seem to presuppose that it is meant to answer a normative question within practical reason. For example, in her discussion of Prichard's response to the question "Why should I do my duty?" Korsgaard assumes that Prichard's answer commits him to the view that moral claims refer to a realm of non-natural, normative properties (1996: 32). But she fails to distinguish this metaethical thesis about the metaphysical commitments of moral judgements from Prichard's normative thesis that moral reasons are foundational or underived. His response to the "why be moral?" question commits him to the latter thesis, not the former. Briefly, his reply is that the question is "improper" or "illegitimate" (Prichard, 2002a: 7, 19), because either it is asking for a self-interested reason to do one's duty, in which case it is seeking the wrong

<sup>8</sup> But in a footnote (14) Korsgaard claims that the evolutionary theory reduces normative ideas to natural ones. This suggests that she is interpreting the evolutionary view to be a reductive account of the meaning of 'moral rightness' or a metaphysical reduction of moral rightness to evolutionary utility, rather than a normative account of the right-making property. Our point is not to suggest that a crude evolutionary account of morality escapes Korsgaard's criticisms, but rather that in her discussion of such an account Korsgaard fails to distinguish the metaethical and normative ethical interpretations of such a position, and this suggests to us that her "normative question" itself blurs metaethical questions of the semantics and metaphysics of moral claims and the normative question of how morality fits into practical reason.

kind of justification of morality since morality's claims are unconditional, or it is seeking a moral reason to do one's duty, in which case it is presupposing the very thing it is asking for. Thus, for Prichard, moral reasons are foundational within practical reason, and do not need to be derived from other practical reasons. Of course, if one holds this position, then, in one sense, placing morality within practical reason will be trivial. This is not to suggest that substantive work will not remain. We have to be convinced that the Prichardian is indeed right about the foundational nature of moral reasons; our brief summary of his position is not a complete presentation of his arguments for this conclusion.<sup>9</sup> Furthermore, the normative ethical task of showing which actions are morally right or morally wrong remains.

There are, no doubt, many objections that one might raise to Prichard's response. Instead of illuminating the status of morality within practical reason, he in the end, one might well conclude, merely dogmatically asserts that it is foundational. However, whatever one thinks of his response to the question "Why be moral?", it expresses a position about the status of morality within practical reason and does not by itself commit him to a position about the semantics, metaphysics, or epistemology of moral judgement. That is, accepting the position that the reasons to do one's moral duty are not derived from any non-moral reasons, but are moral through and through, does not commit one to any thesis about what moral claims mean, what moral predicates such as 'duty' express, or whether and how we come to know moral truths.

Prichard allows that one may come legitimately to doubt whether an action one thought was wrong really is wrong but insists that this is not the same thing as granting that an action is wrong and then wondering whether one has reason to do it (Prichard, 2002a: 18–20). In describing how one resolves such doubt he does apparently commit himself epistemologically. He claims that moral truths "can only be apprehended directly by an act of moral thinking" (Prichard, 2002a: 19). "We do not", he claims, "come to appreciate an obligation by an *argument*, i.e. by a process of non-moral thinking" (Prichard, 2002a: 13). However, to the degree that these genuinely are epistemological commitments, they are detachable from the claim that moral reasons are foundational. A reductive realist, for example, might think that moral reasons are foundational within practical reason, but deny that the epistemology involved is at bottom any different than that of the natural sciences.

Later, Korsgaard says that Prichard's way of asking the normative question, "Is this action really obligatory?" can be understood either as

<sup>9</sup> See in addition Prichard (2002b: 27–9).

asking whether a moral predicate has been correctly applied or as asking the question she is interested in, which is how any obligation can be normative. She claims that this ambiguity led Prichard to mistakenly believe that by showing that the requirement to perform an action can be derived from the principles of the correct moral theory, and thus that the moral predicate 'duty' correctly applies to the action, one has answered whatever request was posed by that form of words (1996: 39). Whether or not Korsgaard is right that Prichard was misled by a failure to distinguish these questions, we claim that her own normative question is itself susceptible to two different interpretations. "How can any moral obligation be normative?" can either express a request to place moral duty within practical reason, a "justification" of morality, or a request for a metaethics of moral judgement—an explanation of what it is to judge that *X* has a moral duty. By failing to distinguish these questions, Korsgaard gives the mistaken impression that, by showing that "substantive realism" is inadequate as an answer to the former question, she has shown that it is inadequate as an account of anything that might reasonably be requested by asking for "the source of normativity".<sup>10</sup>

Korsgaard's discussion of Moore's famous open-question argument in *Sources* is also infected by her failure to distinguish questions within practical reason from metaethical questions. She claims that the open-question argument derives its power from the pressure of "the normative question": "That is, when the concept of good is applied to a natural object, such as pleasure, we can still always ask whether we should really choose or pursue it" (1996: 43). But, she continues, this should not lead us to conclude, as Moore did, that normative concepts do not have criteria of application. Korsgaard seems to think that Moore, like Prichard, failed to distinguish the question whether a normative concept has been correctly applied from "the normative question", and thus that Moore mistakenly thought that because no naturalistic answer can be given to "the normative question", there can be no naturalistic criteria given to guide the application of a normative concept. But, of course, Moore himself claimed that there were synthetic necessary truths connecting normative and natural properties (e.g. pleasure is good)—that is, he would have accepted a naturalistic account of the normative-making properties. There thus is a sense in which he would have accepted that naturalistic criteria can be given

<sup>10</sup> Furthermore, we will argue below that this confusion leads Korsgaard to think that she is giving a full account of the source of normativity, when in fact she is best interpreted as arguing that a certain set of Kantian claims are the most fundamental normative claims of practical reason, not as giving a metaethical account that tells us what those claims mean, what metaphysical commitments we incur by making them, or how we can come to know them.

for the application of normative concepts, although he would have denied that such criteria constitute analytic definitions of normative concepts or that they allow us to reduce normative properties to natural properties.<sup>11</sup> Moore's non-reductive normative realism, although committing him to the claim that the property of good is not identical to any natural property, did not prevent him from accepting a naturalistic account of good-making properties. We will now argue that Korsgaard's attempt in *Sources* to contrast her own view with non-natural realism is spoiled by her failure to notice that non-natural normative realism is compatible with a naturalistic account of normative-making properties.

Korsgaard claims that the obligations that an agent has spring from what that agent's practical identity forbids, where a practical identity is a description of the agent under which he values himself and sees his life as worth living. Thus, for example, if you value yourself as a psychiatrist, you have an obligation not to violate your patient's confidence, since violating a patient's confidence is incompatible with the job description of a psychiatrist (1996: 101). She also claims that the value of an agent's practical identities depends on the value that he places on his own need for practical identities—his humanity (1996: 121). Furthermore, she argues that rational action is impossible unless agents value their humanity, and that therefore human beings are valuable (1996: 124). We do not want to question the truth of any of these claims, although there is much to contest here; rather, we question whether these claims amount to a metaethical position.<sup>12</sup> The problem is that before we can evaluate the metaethical status of such an account, we need to know what it is to *value* oneself under a description. Is this a belief that something, for example, psychiatry, is valuable? If so, then whether such an account is compatible with non-reductive realism all depends upon whether the belief that something is valuable is a belief in a non-natural property.<sup>13</sup>

<sup>11</sup> In fact, elsewhere Korsgaard herself seems to realize this: describing Moore's position, she writes "Of course it might be *true* that the good is pleasure, or the desirable, or what someone wills" (2003*b*: 103). But then it is not true, contrary to what Korsgaard claims in *Sources*, that Moore thinks that there are no naturalistic criteria for the application of normative concepts.

<sup>12</sup> It might be thought that it is the entire transcendental-style argument for the value of humanity, not the premises taken in isolation, which constitutes Korsgaard's alternative metaethical position. We hope that our discussion of Korsgaard's similar transcendental-style argument for the principle of instrumental reason will make it clear why this is not so. But for specific discussion of the metaethical status of Korsgaard's argument for the value of humanity in *Sources*, see Hussain and Shah (2005*a*).

<sup>13</sup> We are not claiming that valuing something is a belief that something is valuable. After all, there are interesting proposals that valuing something is a matter of having a certain hierarchy of pro-attitudes towards that thing. See e.g. Bratman (2000). We

Unfortunately, Korsgaard's commentary on her account does not help us to understand what metaethical position her account is supposed to yield:

In one sense, the account of obligation that I have given in these lectures is naturalistic. It grounds normativity in certain natural—that is, psychological and biological—facts. . . . My account does not depend on the existence of supernatural beings or non-natural facts, and it is consistent with although not part of the Scientific World View. In that sense, it is a form of naturalism. (1996: 160)

The second sentence is vitiated by an ambiguity in the term 'grounds'. If 'grounds' just means "depends upon", then the sentence does not imply the absence of non-natural facts. The fact that something is valuable might depend upon natural facts—for example, that it is pleasurable, or that it is the object of an autonomous choice (if this is a natural fact)—but as long as this dependence is not the strong relation of identity, it is left open whether the fact that *x* is valuable is a non-natural fact about *x*. If Korsgaard were instead using 'grounds' in a non-standard way to mean "identical to", then her account would be a form of naturalistic realism. However, this would conflict with her explicit rejection of the kind of naturalism that "identifies normative truth with factual truth" (1996: 161).<sup>14</sup>

Moore himself would also have thought that once one has determined that, for example, pleasure is good—that is, that the property of pleasure is good-making, not that the concept of pleasure and the concept of good pick out the same property—the question whether one has a reason to pursue pleasure has been answered. He thus would not have understood the open-question argument as targeting the question whether we have a reason

explore non-cognitive interpretations of Korsgaard's account of practical identity in Hussain and Shah (2005*b*). Korsgaard does label her account 'constructivism' to contrast it with non-reductive realism. However, as we shall argue later, it is far from clear whether Korsgaard's characterization of constructivism amounts to a metaethical position. Our point for now is that Korsgaard's account of valuing does not by itself yield a metaethical alternative to realism.

<sup>14</sup> Later in this beguiling passage Korsgaard writes: "From outside that (first-personal perspective) standpoint, we can recognize the fact of value, but we cannot recognize value itself" (161). If value cannot be discerned from the empirical third-person perspective, then naturalistic realism is ruled out. A natural way of interpreting the thought in this quotation is that, from the empirical perspective, we can discern the good-making facts, but it is only from the non-empirical, normative point of view of practical reason that we can see that these facts are good-making, because it is only from such a perspective that we can come to know the normative principles that tell us which natural facts are good-making. While this position implies the denial of naturalistic realism, it is perfectly consistent with non-natural normative realism, which says that normative facts are irreducible to natural facts, but which allows that normative facts are dependent on natural facts; that is, it accepts that the only things that have the property of goodness are natural objects/properties.

to pursue our duty or do what is good. Moore aimed the open-question argument at the *semantic* question whether normative concepts can be defined in terms of natural concepts, and concluded from this argument that they cannot. Furthermore, from this semantic result he inferred that normative concepts refer to irreducible normative properties. But none of these conclusions are answers to the question of the place of moral considerations within practical reason, and thus none answer the normative questions, “Why should I do my duty?” or “Why should I pursue what is good?”

Because Korsgaard fails to distinguish this normative question from the metaethical question that Moore was asking, and misinterprets Moore’s semantic conclusion that good is indefinable (at least in purely naturalistic terms) as attempting to answer the question of whether (naturalistic) criteria can be given to guide the application of normative terms, she fails to come to grips with, much less argue against, Moore’s metaethical non-reductive realism. This is not to say that there is not a legitimate metaethical worry lurking behind Korsgaard’s ill-formed objections to non-reductive realism, which is that Moore’s account does not provide any illumination: we have no account of what it is for a property to be a normative property, and we have no substantive epistemology that explains how we come to know normative facts. Perhaps Moore is right that no such illumination is possible, but one can sympathize with Korsgaard’s inchoate desire for illumination nonetheless.<sup>15</sup> If this is Korsgaard’s dissatisfaction with non-natural, normative realism, she fails to express it correctly, because she fails to disentangle the worry that non-reductive realism fails to illuminate and give a substantive epistemology of the normative properties that it claims are expressed by normative predicates from the worry that non-reductive realism fails to illuminate the place of morality within practical reason.

### 3.2. “The normativity of instrumental reason”

Just as asking for the source of the normativity of duty can be interpreted either as a request to place duty within a conception of practical reason or as a request for a metaethical interpretation of judgements such as “One has the duty to provide for one’s children”, so too asking for the normative foundation of the principle of instrumental reason can either be interpreted as a request to place the principle of instrumental reason within a conception of practical reason or as a request for a

<sup>15</sup> Another worry that Korsgaard might be trying to express is that the non-natural normative realist cannot explain how moral facts are able to motivate us. Below, we discuss this interpretation of Korsgaard’s worry in connection with her discussion of the realist position about the principle of instrumental reason.

metaethical interpretation of the principle (or of particular means–ends normative judgements). In her article, “The Normativity of Instrumental Reason”, Korsgaard certainly, once again, can come across as taking a metaethical position. She introduces “the Kantian conception of practical rationality” as a “third and distinct alternative” to be distinguished from, and preferred to, “empiricist” accounts, on the one hand, and “realist” or “dogmatic rationalist” positions on the other (Korsgaard, 1997: 219). Such metaphysical and epistemic sounding labels are once again hard not to read as labels for positions identified by metaethical commitments. And again her setting of her position in contrast to positions with such labels suggests that she takes herself as defending a distinctive metaethical position.

We will approach the question of whether she is indeed expressing a metaethical position in this article in two stages. First, we will consider her arguments against what she calls the realist position. As in the moral case, her failure to distinguish between different questions about normativity confuses the issue here. On the normative reading of her objection to the realist, the realist can deploy a Prichard-style response. Such a style of response shows that the realist position could be coherent. More importantly, as in the moral case, this response has nothing in particular to do with any metaethical position. If we try to read her worry about the realist in metaethical terms, then it is much harder to see what her objection is supposed to be—though we will consider a couple of possibilities. In any case we assume that whatever metaethical objections can be read out of her discussion are supposed to be objections to which her own view of instrumental reason is immune. The discussion in this section will thus set the stage for a consideration in the next section of the apparent positive position expressed in the article—the Kantian conception of practical reason according to which the principle of instrumental reason is constitutive of the will.

#### 3.2.1. *The critique of non-reductive realism*

She identifies the realist position initially as the view that moral requirements must have “some sort of ontological foundation, by positing the existence of certain normative facts or entities to which moral requirements somehow refer” (Korsgaard, 1997: 218). This is the view adopted by the “dogmatic rationalists”, a term that she then uses interchangeably with “realist” for the rest of the article.<sup>16</sup> Given that the epistemic-sounding “dogmatic

<sup>16</sup> We note immediately one potential source of confusion here that arises from mixing ontological and epistemic labels. An empiricist can normally, though not apparently in Korsgaard’s idiosyncratic terminology, be a realist. Of course the empiricist conception of the ontology will be such as to fit his epistemology—normative facts must be knowable by empirical means.

rationalism" is the alternative label for what she calls realism, we take it that Korsgaard intends this label to express the position that the relevant normative facts are not empirically accessible and so presumably are also not naturalistic or material facts. The epistemology of this position is to be some kind of intuitionism.<sup>17</sup>

She claims that the difficulty for non-reductive realism "exists right on its surface, for the account invites the question why it is necessary to act in accordance with those reasons, and so seems to leave us in need of a reason to be rational" (1997: 240). As we think is clear from the context, Korsgaard must mean "rationally necessary" as opposed to, say, metaphysically or causally necessary. We also assume that the point here is *not* that, though there can be reasons to  $\phi$ , I perhaps ought not to  $\phi$  because the balance of reasons favours not  $\phi$ -ing. Putting the point more clearly in terms of an ought, then, the question supposedly invited is "Why ought I to do what I ought to do?" The Prichard point is that such a question does not make sense. If I have accepted that I ought to  $\phi$ , then how can it still make sense for me to ask why it is rationally necessary to  $\phi$ ?<sup>18</sup> To think that I ought to  $\phi$  just is to think that it is rationally necessary to  $\phi$ . The point has nothing to do with the metaethical issues of what kind of mental state I am in when I think that I ought to  $\phi$  or whether there is a mind-independent fact accessible only by rational intuition that I ought to  $\phi$ .

Thus understood as one kind of normative question it is hard to see what the objection is. Of course, there are questions we can sensibly ask our Prichardian about instrumental reason, including further normative questions. We can ask what makes particular considerations reasons—the reason-making features. The Prichardian might respond in predictable ways.<sup>19</sup> What makes the fact that  $\psi$ -ing is a means to  $\phi$  a reason to  $\psi$  is that you, say, desire to  $\phi$ .<sup>20</sup> We are not sure about this, but Korsgaard seems to suggest that this would be to derive an "ought" from an "is" (1997: 245).

<sup>17</sup> When she turns to specifying what such a realism about instrumental rationality would look like, we get an additional ontological claim: "truths about reasons . . . exist independently of the will" (219). Finally, when she turns in earnest to the discussion of realism she says that according to realism "there are facts, which exist independently of the person's mind, about what there is reason to do" (240). Note that independence from the will is not identical to independence from the mind. We mention all of this to emphasize how, in one way, the so-called realist target is quite limited—it is a subset of realist positions out there—and how it is very clearly specified in terms of epistemological and metaphysical features.

<sup>18</sup> Cf. Dreier (2001).

<sup>19</sup> Not Prichard himself since he was quite suspicious of the idea that there might be, as he would put it, such "general" answers available (Prichard, 2002c: 62; 2002a).

<sup>20</sup> We do not endorse this version of the principle of instrumental reason. There are several issues here. (i) How should the antecedent be specified? Should it be specified in

But one is no more deriving an "ought" from an "is" in this case than when one says that you should mow the lawn because the grass is tall or that what happened to her is bad because she is in pain. Non-normative facts will *make* certain normative claims true in any non-error-theoretic account. This is just a result of the fact that normative properties rarely, if ever, apply barely.<sup>21</sup>

Our Prichardian might well grant that there is a general normative truth in the background expressible by some version of the following:

- (1) For all  $S$ ,  $\phi$  and  $\psi$ , if  $S$  (believes that  $S$ ) desires/intends that  $S$   $\phi$  and ( $S$  believes that)  $S$ 's  $\psi$ -ing is a means to  $\phi$ , then  $S$  (believes that  $S$ ) ought to/has reason to  $\psi$ .

For all we have said the Prichardian can think of this as another premise that the agent *believes* and then combines in his reasoning with the following beliefs:

- (2) My  $a$ -ing is a means to  $b$ -ing
- (3) I desire to  $b$

to reach the conclusion

- (4) I ought to  $a$

The inference principles relied on are just the ones of theoretical reason.

The alternative, which is also open to the Prichardian, would be to introduce the instrumental principle as a practical inference principle in its own right. He would also add that what makes following that principle correct is precisely that the associated normative claim is true.

In "Realism and Constructivism" Korsgaard argues that the "realist account of the normativity of the instrumental principle is incoherent"

terms of desiring, intending, or willing? Should the antecedent be normative? (ii) Should the principle allow for detachment? (iii) Should it should be a "strict" or a "slack" demand? See Broome (2000) for a discussion of (ii) and (iii). For an extended discussion of the principle of instrumental reason, see Stalzer (2004).

<sup>21</sup> She makes a similar mistake in her discussion of Derek Parfit. She seems to think that Parfit—whom we can treat as a non-reductive realist—has to choose between two views: first, that the complex considerations we use in determining whether an action is right will not explain why the action is right—"It is right because it has the property of rightness"; second, that these considerations "constitute its rightness" (Korsgaard, 2003a: 3). Now Parfit should answer that neither is the case. The considerations *make* an action right but do not *constitute* rightness. When the considerations are complicated, then it might well be hard work to come to know that the action in question is right. Indeed an action is right because it has the property of rightness, but we can still ask why it has the property of rightness and this leads us to the considerations that make it right. Talk of constitution though threatens, unless the possibility is explicitly ruled out, to be a reductive view and Parfit wants to avoid that.

(Korsgaard, 2003b: 110). The instrumental principle would have to be “some sort of eternal normative verity”. “How”, she asks, “is this verity supposed to motivate him?” (110). The picture, she claims, is incoherent: “The point is that the instrumental principle cannot be a normative truth that we *apply* in practice, because it . . . is essentially the principle of application itself, that is, it is the principle in accordance with which we are operating when we apply truths in practice” (110). Applying the normative truth, the principle of instrumental reason, presupposes that we are already applying the principle.<sup>22</sup>

The point, though, does not have to be put in terms of the principle of instrumental reason. Consider the following more general reconstruction of Korsgaard’s point here. The heart of the internalist thought is that normative beliefs are practical. That is, in order to have any normative belief I must be able to act on it. Put in the most general terms, the agent has to be applying or following something like the following normative requirement in order to have any normative beliefs:

$$(5) O(BO\phi \rightarrow \phi)$$

The symbolism here is basically Broome’s: believing that you ought to  $\phi$  requires  $\phi$ -ing. We are proposing to read  $\phi$  generously to allow, for example,  $O\phi$  to be a statement of the general instrumental principle like (1): roughly, believing that you ought to be instrumentally rational requires being instrumentally rational.

Now Korsgaard’s point is that an agent cannot be motivated by a belief with the content (5) unless he is already applying (5). Imagine giving the agent the following belief

$$(6) BO(BO\phi \rightarrow \phi)$$

This is just another belief of the form  $BO\phi$ . No consequences for motivation follow unless the agent is already applying (5). And so perhaps it is not even possible for the agent to have a belief with normative content without already applying (5).

We agree that something like this seems right and similar points will hold for some theoretical norms. It may be true that one does not count as having beliefs unless one is thinking correctly to some extent. If so, then one cannot come to apply a fundamental principle of thinking on the basis of believing

<sup>22</sup> Compare her comment about “goodness in action” earlier on p. 110: “To put the same point another way, goodness in action cannot just be a matter of applying our knowledge of the good—not even a matter of applying our knowledge of what makes action itself good. This is because the ability to apply knowledge *presupposes* the ability to act.”

a normative claim, since one would already need to apply the principle in order to believe the normative claim. If this is Korsgaard’s point, the realist should happily grant it, as it is perfectly compatible with his central claim that the truth of the normative claim makes it correct to follow the principle. A normative truth can make a certain way of thinking correct even if it cannot be an agent’s grounds for coming to think that way. This is just the point that what makes something correct to do and what one’s grounds are for doing it need not coincide.<sup>23</sup> In fact, the realist can go further and claim that an agent can come to believe that the normative fact that he ought to be instrumentally rational makes it correct to follow the instrumental principle—which is the principle that he follows in arriving at this very belief—even if the fact that he ought to be instrumentally rational cannot be his initial ground for following the instrumental principle. Thus while Korsgaard may have shown us that normative facts cannot play a certain epistemological role, she has not shown us anything that the non-reductive realist cannot take in his stride.

### 3.2.2. Interpreting the positive account

To bring out further how Korsgaard’s criticism of non-reductive realism misses the mark, and more importantly, to show that Korsgaard does not in fact commit herself to a metaethical position, we will now argue that, for all Korsgaard says in “The Normativity of Instrumental Reason”, non-reductive practical realism is compatible with her own “explanation” of the normativity of instrumental reason. This is because what she says only commits her to an explanation of the place of the principle of instrumental reason within practical reason, not to a metaethical interpretation of what it is to think normative thoughts such as that one ought to take the necessary means to one’s end.

In trying to reconstruct a positive metaethical position from “The Normativity of Instrumental Reason”, we will end up considering several possibilities for such a reconstruction; however, we begin with what naturally comes across as a family of potentially distinctive metaethical claims, namely, the claims that the will or action are supposedly constituted by certain principles or norms.

Now once again her introduction of her position seems to be driven by the normative question about the place of the instrumental principle within practical reason: “The [realist] model, as I said earlier, seems to invite the question: but suppose I don’t care about being rational? What then? And in

<sup>23</sup> Thanks to David Velleman for drawing our attention to the relevance of this distinction.

Kant's philosophy this question should be impossible to ask" (1997: 244). It is not clear, as we have emphasized, what is meant by "care". If what Korsgaard is trying to get at is a possibility where someone accepts that  $\phi$ -ing is rational, but then proceeds to ask whether it is rational to  $\phi$ , then the realist can insist that the question does not make sense though, as we have emphasized, the realist's insistence is independent of his metaethical position—his realism.

But what is Korsgaard's central positive claim? "To will an end just is to will to cause or realize the end, hence to will to take the means to the end. This is the sense in which the principle is analytic. The instrumental principle is *constitutive* of an act of the will. If you do not follow it, you are not willing the end at all" (244). The problem, as Korsgaard realizes, is that this does not seem to allow for the possibility of instrumental irrationality. If it is logically impossible to will an end without taking the means to the end, then it is impossible to be instrumentally irrational—to will an end and fail to take what one recognizes to be the means to that end.

To prevent this she makes one negative claim about willing: "So willing the end is neither *the same as* being actually disposed to take the means nor as being a particular mental state or performing a mental act which is *distinct from* willing the means" (1997: 245). And a positive claim about willing:

[W]illing an end just is *committing* yourself to realizing the end. Willing an end, in other words, is an essentially first-personal and normative act. To will an end is to give oneself a law, hence, to govern oneself. That law is not the instrumental principle; it is some law of the form: Realize this end. That of course is equivalent to 'Take the means to this end'. So willing an end is equivalent to committing yourself, first-personally, to taking the means to that end. (1997: 245)

There is a lot packed in here that one wishes had been laid out a bit more slowly. Our hope is to develop the different possible charitable interpretations of Korsgaard's position and see how far they go.

Put aside for a moment that talk of "laws". Now, much of this actually sounds like many a Prichardian of our acquaintance—though we realize perhaps not yours. To allow for differences to emerge, let us use a term other than "willing" for the kind of attitude or act that the Prichardian wants to claim is directed at the end in the cases in which the principle of instrumental reason applies: we will use the term "intending". So our Prichardian states a string of conceptual truths:

Intending an end just is *committing* yourself to realizing the end. Realizing the end requires taking the means to the end. So committing yourself to realizing the end is equivalent to committing yourself to realizing the means to your end. *Following* the instrumental principle is committing yourself to realizing the means

when you commit yourself to realizing the end. So you don't count as committing yourself to realizing the end unless you are following the instrumental principle. The instrumental principle is *constitutive* of intending. If you do not *follow* it, you are not intending the end at all. And, of course, if you're not following it, you're also not being rational.

The plausibility of these conceptual truths turns no doubt on taking "committing", "following", and "intending" as normative concepts. Is this a problem for the Prichardian? It is not at all clear why it would be. The Prichardian, whatever his metaethics, is no error theorist. He is happy to make claims involving unanalyzed normative concepts and he is convinced, let us imagine, by Korsgaard that these sound like good ones to make. But if he can make them too, then these claims seem, unsurprisingly to us, to be merely normative claims and not part of any distinctive kind of metaethical position.<sup>24</sup>

But what about the claim that "willing the end is neither *the same as* being actually disposed to take the means nor as being a particular mental state or performing a mental act which is *distinct from* willing the means" (Korsgaard, 1997: 245). Well, our Prichardian could happily go along with Korsgaard here, but could also deny the claim. He *could* insist that the mental act of willing the end is distinct from the mental act of willing the means. Korsgaard has the odd view that the "dogmatic rationalist conceives willing an end as being in a peculiar mental state or performing a mental act which somehow logically necessitates you to be in another mental state or perform another mental act, namely, willing the means" (1997: 244). But it is hard to imagine our Prichardian saying anything like that for precisely the reason Korsgaard goes on to give: "for no mental state can logically necessitate you to be in *another* mental state or perform another mental act" (1997: 245). The Prichardian would claim that being in one state *rationally, not logically*, necessitates being in the other state or performing the other mental act. So there can be two distinct states or mental acts. It is just that being in one involves a commitment to being in the other.

But what about the Prichardian who is committed to a realist metaethics and a non-reductive one at that? Didn't we grant that Korsgaard might well have legitimate worries about this view? And surely there must be something in Korsgaard's view that is meant to be able to respond, or avoid, precisely those worries that then will also allow us to distinguish her position from the realists. Or at least her claim that the will is constituted by normative principles adds something metaethically distinctive, whatever metaethical label we want to apply to the resulting position. As it turns

<sup>24</sup> There will be particular metaethical positions that might rule them out.

out, nothing Korsgaard says implies anything of the sort. Consider first the epistemological worry that the realist has no account of how we come to know normative facts. As we have just noted though, the normative claims the Prichardian has just endorsed are all apparently conceptual claims. Issues remain. First, arguably the Prichardian will have to step out of the circle of conceptual claims at some point. In the case of instrumental reason, surely at some point the agent will have to will an end. Can the thought that I am willing an end be a conceptual claim? Implausible. So appealing to some supposedly unproblematic epistemology for conceptual claims will not resolve the epistemological issue. The realist thus seems to be left without an illuminating account of how I can come to know that I am willing an end, at least if he takes this to be a normative claim. If I could know that I am willing, then the conceptual claims listed above, plus a means–end belief, should take me the rest of the way—not through action, but to a commitment to taking the means, to willing the means.

But has Korsgaard made progress on the epistemic front? Maybe she thinks that because this kind of normative fact is not mind- or will-independent, there is no need for a substantive tracking epistemology. But we need to be very careful. Again the distinction between normative-making properties and the normative property itself is crucial. To say that I am a bad person because I want to sleep with my neighbour's wife is not immediately to claim that this normative fact is mind-dependent just because the desire is a mental state of mine. Or if it is sufficient to make the normative fact mind-dependent, then there is no “realist” opponent. Wrong-making properties can be mind- and will-dependent and many, if not most of them, are. What is important is whether the normative property is mind-dependent. Merely saying that willing is normative does not make the normative property mind-dependent or will-dependent in any interesting sense. Wrongness is not desire-dependent because my having a certain desire can be wrong. So even if willing is normative, the normative property could be mind-independent. And if it is mind-independent, then it is not obvious how I come to know that it is instantiated.

So perhaps we should take Korsgaard as suggesting that the normative property is not mind-independent. Now, if all Korsgaard adds to this is the brute claim that the normative property *is* mind-dependent then we do not really have a distinctive metaethics, nor one that really provides any help on the epistemic front. As far as the brute claim of mind-dependence goes, our Prichardian can happily keep step—normative properties are mind-dependent but no reason has been given yet to think of them as any less real. No reason not to either of course. The problem is that a brute mind-dependence claim does not get us very far. For all we have said, the view could be the magical one: when I am in a certain mental state the

normative property itself somehow comes into existence and not just its particular instantiation. Such metaphysical dependency does not yet imply any particular story of epistemic access. There are complicated issues here, but the main point is that simply declaring mind- or will-dependence does not really get us anywhere.

The problem we may seem to have been circling around though is the question of what willing itself is. Is it something mental? Surely it must be. For Korsgaard willing is “essentially” normative (1997: 245). Perhaps here lies the distinctive sense in which the normative is will- and mind-dependent and perhaps here lies the solution to our epistemological worries.

What does it mean to say that something is essentially normative? Can the “realist” endorse such claims? Well, the realist does make claims such as that murder is essentially wrong. But perhaps this is confused because the realist has to identify something that has the normative property. It is the particular act of killing that has the normative property of being murder. This way of putting it suggests that one can distinguish between the thing of which the normative property is being predicated and the normative property itself. And this might suggest either that we can identify the entity (here an act) without using normative language or, in fact, that it is not essential to the act that it have this normative property—the normative property is not an essential property of the act in question. Perhaps then Korsgaard's distinctive suggestion would be that the act, for her the act of willing, is essentially normative in the sense that it cannot be identified without using the normative language and so the normative property is essential to it in the way that it is not to the act of Jill killing John.

However, there is no reason for a non-reductive realist to grant all this. There may well be acts of killing that are not murder, but the realist can claim that *this* act of Jill killing John *is* murder and essentially so. And he may well deny that there is any way to identify it in terms other than the normative. Similarly, without the normative concepts of belief, desire, intention, we cannot identify an action. A sequence of events standing in causal relations identified without these terms may not, such a view might insist, line up with a description in terms of beliefs, desires, intentions, and actions.<sup>25</sup> This is the sense in which our non-reductive realist's picture could involve a commitment that might capture what Korsgaard is getting at in her talk of the “first-personal”. There is no identifying actions in nomological vocabulary. Of course this does not mean that I can only make judgements about myself. Most of our moral practices rely on our ability to make judgements about when others have acted and what they intended.

<sup>25</sup> Cf. Hornsby (1997: 295–6).

We would basically have to give up our existing moral practices if we could not make such judgements—if it literally only made sense for me to judge about willing in the first person.

So there does not look like anything in the talk of willing being essentially normative that the realist has to deny. Now there are puzzles here for the realist, but these, as we shall try to show, are just as much puzzles for Korsgaard.

There is the question of how the level of normative vocabulary “fits” with the level of physical or nomological vocabulary. This is a classic puzzle of course central both to Kant, those inspired by Kant, and to contemporary discussions in the philosophy of action and mind. We do not intend to defend the claim that there is really a difficult problem here or that non-reductive realist approaches will not work. Our point is just that, as far as we can see, Korsgaard says nothing here that contributes to the debate. What she does say could, as we have suggested, be interpreted in non-reductive realist terms. Perhaps this interpretation would lead to a solution—consider the writings of McDowell, Hornsby, or Dancy, or for that matter Davidson’s own view. The point though is that Korsgaard has not provided the Kantian with any distinctive way of solving this problem.

Such a non-reductive realist also may have a puzzle when it comes to the question of concept acquisition. How is it that I acquire these normative concepts in order even to have these thoughts? This question is closely related to the worry that the non-reductive realist has no non-trivial account of the content of the normative thoughts. Now, whether concept acquisition is a problem will depend on what answer we give to this question and various background considerations, but again there is nothing Korsgaard says that will help with any of this.<sup>26</sup>

Korsgaard gives no non-trivial account of the content of the relevant normative claims and thoughts beyond the conceptual claims mentioned already. Defending this claim does allow us to fill in a lacuna in our discussion of Korsgaard. We had put aside the talk of laws and legislation right at the beginning of our attempt to give an interpretation of Korsgaard’s positive view. It is therefore true that she does say more about willing than we have considered above.

<sup>26</sup> As long as one is not concerned about an error theory or some other kind of confusion lurking in the conceptual scheme, then giving no account of the content is probably fine for practical purposes. We can happily go on using our concepts without any such metaethical account. But this is again a point that does not differentiate between Korsgaard and the non-reductive realist.

To will an end is to give oneself a law, hence, to govern oneself. That law is not the instrumental principle; it is some law of the form: Realize this end. That of course is equivalent to “Take the means to this end”. So willing an end is equivalent to committing yourself, first-personally, to taking the means to that end.

We are not quite sure how equivalence relations work between laws, but we are happy to grant that perhaps it is a normative truth—conceptual or not—that the law of the form “Realize this end” is equivalent, in some sense, to a law of the form “Take the means to this end”. We are even happy to grant that the command “Realize this end!” is the same as the command “Take the means to this end”—not that we have any special account of the content of commands to provide here. The problem though is that the language of giving oneself a law and governing oneself is surely normative language and so by itself does not help with the metaethics or with providing us with a non-trivial account of the content.

Here is one way to see the point. What is the difference between intending an end and giving oneself a law of the form “Realize this end!”? Korsgaard will want to insist that there is no difference but, as always with such conceptual claims, the account of intending in terms of self-legislation will only be illuminating if we have some understanding of what it is to self-legislate *other* than just intending an end. The problem is that when we shift to some attempt to elucidate what it is to give oneself a law we either have no further elucidation or we end up relying on non-normative reductions that seem quite implausible. So consider how Korsgaard elaborates on the talk of self-legislation: “Then what does it mean to say I take the act of my own will to be normative? Who makes a law for whom? The answer in the case of the instrumental principle is that I make a law *for me*. And this is a law which I am capable of obeying or disobeying” (246). What is required is “that there be two parts of me, one that is my governing self, my will, and one that must be governed, and is capable of resisting my will” (247–8). Does any of this help to elucidate the content of the normative claims? Not really. The problem is that the language being used is both metaphorical and still normative. To the degree that we have a grasp of what it is to make laws, what it is to “give” someone a law and what it is to govern someone, it is because we are competent users of these notions when we talk about kings or legislatures—political sovereigns—issuing laws that their subjects are to follow. In their normal context of usage the concepts here are themselves normative and in philosophical discussions of them their metaethics are contested.<sup>27</sup> The problem is only compounded by the fact that the natural way to read their deployment in Korsgaard’s work is surely as metaphor.

<sup>27</sup> As we well know from discussions in political philosophy, it is not at all easy to say what political authority or sovereignty comes to or what the nature of political obligation

There is no political state, no court system, no police *literally* in the self, though there is a long tradition of talking metaphorically as though there were rulers and ruled and so on. It is hard to see how the metaphor says anything more than simply that I can will an end, but fail to will the means even though I should.

Korsgaard's "explanation" of normativity in the end, then, merely uses all of the normative notions that we were hoping for an elucidation of and thus fails to constitute an account of the content of normative claims. This is fine, as we keep emphasizing, if she is doing normative ethics, but of course it prevents her from presenting any metaethical position. Thus whatever worries Korsgaard might have about the non-reductive realist's inability to say more about "*what it is* that [an agent] recognizes" when he recognizes "certain considerations *as* reasons"—to say, in other words, more about normative content—the same worries apply to her view. Indeed the unwillingness to step outside the circle of normative concepts is strikingly similar, however legitimate it might well be. The non-reductive realist at least adds some claims about the nature of normative facts, the semantics of normative language, and the nature of the attitudes that we have towards normative content—he at least says things that look as though they will commit him metaethically. Korsgaard seems to reject these additions of the non-reductive realist. Her remaining claims though are completely compatible with what a non-reductive realist would say. She does not replace the realist's positive metaethical claims with any of her own and so there appears to be no new metaethical position expressed.

Of course it would be a mistake simply to conclude from all of this that Korsgaard and the non-reductive realist have the same metaethical position. The fundamental problem is that her claims are compatible with different metaethical positions or interpretations.

#### 4. Constructivism

Korsgaard claims the banner of constructivism for her view, and she clearly sees constructivism as an alternative to non-reductive realism. We will finish by considering whether her thoughts on constructivism amount to a distinctive metaethical position.

Korsgaard suggests in "Realism and Constructivism" that the difference between the cognitivist position she labels "substantive realism" and her own

is. And as we well know from the discussions in the philosophy of law between legal positivists, natural law theorists, and legal realists, it is not at all clear what indeed calling something a law comes to.

"constructivist" alternative lies in their views of the *function* of normative concepts: The substantive realist thinks that the function of normative concepts is to describe normative reality, whereas the constructivist thinks that the function of normative concepts is to label the solutions to practical problems of what to do (Korsgaard, 2003*b*: 116). But this description of the function of concepts does not yet reveal whether Korsgaard has an alternative metaethics, nor does it set up a contrast with the substantive realist's traditional conception of normative concepts. First of all, if the problems themselves are couched in normative terms, Korsgaard's description of the function of normative concepts as labelling solutions to these problems will not help to establish a metaethical position. The problem is that, in order to understand this function, we would first need to understand the normative concepts that express it. So we would already need an account of normativity before we could use Korsgaard's account to grasp the functional distinction she wants to draw. Second, why should we think that the function of describing normative reality and the function of solving practical problems of what to do conflict? Surely the substantive realist will agree that ethics is about finding the solution to the problem of what to do, adding that normative facts (e.g. action *A* has the property of to-be-doneness) provide the answers to these questions. Correctly describing normative reality, discovering which actions have the property of to-be-doneness, answers the question of what one should do.

Korsgaard does say more about constructivism, or as she also calls her position, procedural realism, in *Sources*.<sup>28</sup> So we will examine her discussion there to see if anything she says establishes a genuine alternative to non-reductive realism. Korsgaard argues that her own view is a form of procedural realism, as opposed to substantive realism, which she is arguing against. Here is Korsgaard's initial characterization of the difference:

Procedural moral realism is the view that there are answers to moral questions; that is, that there are right and wrong ways to answer them. Substantive moral realism is the view that there are answers to moral questions *because* there are moral facts or truths, which those questions ask *about*. (1996: 35)

But this way of putting things makes substantive realism out to be a species of procedural realism, as the substantive realist agrees with the procedural realist that there are answers to moral questions, and that there are right and wrong ways to answer them. The substantive realist also gives a particular

<sup>28</sup> Although we will argue below that, given the way that Korsgaard uses the terms, procedural realism and constructivism are not equivalent positions. Procedural realism is a broader category of which substantive realism and constructivism are meant to be distinct species.

explanation of what answers moral questions, namely moral facts, but this is not excluded by the description of procedural realism given above. That is, nothing in the above specification of procedural realism excludes the claim that there are procedure-, mind-, or will-independent moral facts. Therefore the specification fails to set up a contrast between procedural realism and substantive realism.

Korsgaard's point, we take it, is that procedural moral realism does not force the acceptance of substantive realism, and might be filled out in a way that makes no commitments to mind-independent intrinsically normative entities. Specifically, it allows that the answers to normative questions are "the results of some constructive procedure" (1996: 35). But of course even a substantive realist can allow that we need procedures for arriving at true moral beliefs, since these procedures are what allow us to track the moral facts. Thus, talk of moral answers being the result of a procedure does not by itself establish the needed contrast between substantive realism and the type of realism Korsgaard wants to advocate.

Does it help to be told that the procedure is "constructive"? Well, what does this mean? It sounds as if it means that the employment of certain procedures creates a normative entity. Thus rather than saying correct procedures track independently existing moral facts, the constructivist claims that the moral facts are created by the employment of these procedures. As Korsgaard puts it:

The procedural moral realist thinks that there are answers to moral questions *because* there are correct procedures for arriving at them. But the substantive moral realist thinks that there are correct procedures for answering moral questions *because* there are moral truths or facts that exist independently of those procedures, and which those procedures track. (1996: 36–7)

Strictly speaking, what she says here does not entail that procedural moral realism is committed to the claim that moral facts are created by the employment of correct procedures, but it is difficult to understand what else could make sense of the non-tracking relation between the correct procedures and moral facts that she has in mind. In any case, it is fairly clear from Korsgaard's overall position that the relation between correct procedures and moral facts that she intends is one, in some sense, of "construction" or creation.

Do we now have the needed contrast between substantive and procedural realism, or better, the contrast between substantive realism and constructivism, since as we noted "procedural realism" seems to denote the category of which substantive realism and constructivism are species? We cannot answer this question until we have a fuller specification of constructivism. First of all, we need a characterization of the procedures that construct the

moral facts. The key question here is whether the specification of these procedures employs normative concepts. For example, Rawls's specification of the procedures that are used to construct justice makes use of the concept of the reasonable. Again, there is nothing wrong with this so long as one is engaged in giving an account of justice within normative ethics, as Rawls is. But this will not do if one is attempting to give a metaethical account of the concept of justice, as it leaves us with the unexplained normative concept of the reasonable. And if we try to give a constructivist account of the concept of the reasonable, we face a regress so long as the procedures used to construct the normative facts expressed using this concept are themselves couched in normative terms.

What if the metaethical constructivist tries to avoid this problem by specifying the correct procedures in non-normative terms? Do we now have a full-blown metaethical position? The problem is that, even if the constructivist is able to specify the relevant procedures in non-normative terms, he is committed to the normative judgement that these are the *correct* procedures. After all, it is only if the procedures are the correct ones that one can use them to construct the relevant normative facts; presumably not any old procedure has the power to create normative facts. But then, what is the constructivist's metaethical account of what it is to judge that a procedure is correct? At this point, all the familiar realist metaethical options appear open. For example, it is open to the reductive realist to say that the relevant normative property (e.g. justice) is identical to the property of being the output of these procedures, specified in non-normative terms. Since a social institution's property of being just and its property of being the output of these procedures are identical, the question of the correctness of the procedure does not come up.<sup>29</sup> But it is also open to the non-reductive realist to say that the procedure has the separate normative property of correctness. Might the constructivist try to give a constructivist account of the judgement that a procedure is correct, going constructivist all the way down, as it were? It is hard to see how this would really provide much elucidation. If the constructivist says, for example, that a certain procedure, call it *x*, creates the facts about which procedure of justice is correct, he must claim that *x* is the correct procedure for constructing facts about correct procedures of justice. And then we are left with an unexplained normative concept, since we still need to be told what it is to judge that procedure *x* is correct.

Might the constructivist specify a procedure that creates its own correctness? It is hard to know what this would involve. If it just means that the

<sup>29</sup> Note that the reductive naturalist could propose this property identity either as an analytic truth or as part of an explanatory empirical theory of the relevant normative discourse.

procedure is self-validating, then it will not single out a unique procedure, since there are many procedures that are self-validating. And then the question arises as to which of these procedures is the correct one. Even if this problem could be avoided, we still would not have an account of what it is to make a correctness judgement, we would just have an account of which correctness judgements are correct. Put it this way: we ask what it is to make a normative judgement. We are told that normative judgements are about normative facts that are created by correct procedures. We then ask what it is to call a procedure correct, and we are told that correct procedures are those procedures that are constructed by a certain procedure,  $x$ , which is the correct procedure for creating correct procedures. We then ask what it is to judge that  $x$  is correct. And now what? We never seem to get outside of the normative circle. The objection here is not that the constructivist fails to give us a non-normative reduction of normativity, since we are allowing for positions such as non-reductive realism that commit themselves to sui generis normative properties. The problem is that non-reductive realism or some other metaethical position needs to be added to constructivism in order to turn it from an account of which normative judgements to make into an account of what it is to make a normative judgement. But if this is so, then constructivism is not really a metaethical position at all.

There is another question that needs to be answered before we can assess constructivism as a full-blown metaethical alternative, even if the constructivist can somehow avoid the problem we just laid out: what does it mean to say that the normative facts are *created* by correct procedures? It is important to recall here the point about right-making properties. Imagine that one of our students follows his usual complicated "procedure" for cheating on an exam. That this procedure was followed will make it the case that he has done something bad. In one sense then the normative fact was *created* by the procedure. Emphasizing "created" here though sounds rather dramatic and a non-reductive realist will insist that our student has merely ensured that the relevant normative property is instantiated. He did not create the property itself and indeed the "creation" of the normative fact was not something that, in one important sense, he could control: once he had followed his "procedure", the badness of his action was not up to him. Thus if normative facts were being created by procedures in only this sense, we would not have an alternative to substantive realism. However, when we try to imagine what it would be for normative facts to be created in some more substantive sense, then the view does start to sound quite magical. The mysteriousness of such acts of creation *ex nihilo* seems to be on the same order as the ontological and epistemic mysteries of the non-reductive realist's intrinsically normative entities and our supposed intuitive access to them. Finally, for all that has been said, once these normative facts are

created they could be just like the substantive realist's normative facts. It is not obvious here that different ontology entails different ontology. Much more clarification would be needed before we could know how different such a view really would be from substantive realism and what kinds of theoretical costs such differences would incur.

## 5. Conclusion

How is it, then, that Korsgaard has put herself in the awkward dialectical position of framing the Kantian position in opposition to non-reductive realism and yet presenting us with a position whose content is compatible with it? As we have argued throughout the paper, the fundamental problem occurs at the very beginning. In framing her inquiries about the "source of normativity", Korsgaard fails to distinguish the metaethical question of what it is to make a normative judgement from normative questions about which normative judgements to make or even which normative judgements we cannot help but make. This leads her to mistakenly think, for example, that by making a strong case for a Kantian position that certain normative judgements are constitutive of agency, she has given an alternative to the non-reductive normative realist's position about the meaning, metaphysics, and epistemology of these normative claims. Her opposition to non-reductive realism similarly suffers from a misunderstanding of the metaethical tasks this theory seeks to accomplish. This causes her to take them to task for failing to give plausible answers to questions of normative ethics, when in fact, for all she says, her own answers to these questions are available to non-reductive realists. While her positive claims constitute a Kantian position on foundational questions in practical reason, they do not constitute a metaethical position—they do not constitute an alternative to non-reductive realism.

## REFERENCES

- Bratman, Michael, 'Valuing and the Will', *Philosophical Perspectives*, 14 (2000), 249–65.
- Broome, John, 'Normative Requirements', in J. Dancy (ed.), *Normativity* (Oxford: Blackwell Publishers, 2000), 78–99.
- Dreier, James, 'Humean Doubts about Categorical Imperatives', in E. Millgram (ed.), *Varieties of Practical Reasoning* (Cambridge, Mass., and London: MIT Press, 2001), 27–48.
- Hornsby, Jennifer, 'Agency and Causal Explanation', in A. R. Mele (ed.), *The Philosophy of Action* (Oxford and New York: Oxford University Press, 1997), 283–307.

- Hussain, Nadeem J. Z., and Shah, Nishi, 'Is Constructivism an Alternative to Realism?' MS (2005a).
- 'Metaethics and its Discontents: A Case Study of Korsgaard', MS (2005b).
- Korsgaard, Christine, *The Sources of Normativity* (Cambridge: Cambridge University Press, 1996).
- 'The Normativity of Instrumental Reason', in G. Cullity and B. Gaut (eds.), *Ethics and Practical Reason* (Oxford: Oxford University Press, 1997), 215–54.
- 'Normativity, Necessity, and the Synthetic a priori: A Response to Derek Parfit', MS (2003a).
- 'Realism and Constructivism in Twentieth-Century Moral Philosophy', *Journal of Philosophical Research*, APA Centennial Supplement (2003b), 99–122.
- Prichard, H. A., 'Does Moral Philosophy Rest on a Mistake?' in *Moral Writings*, ed. J. MacAdam (Oxford and New York: Clarendon Press, 2002a), 7–20.
- 'Duty and Interest', in *Moral Writings*, ed. J. MacAdam (Oxford and New York: Clarendon Press, 2002b), 21–49.
- 'Kant's *Fundamental Principles of the Metaphysics of Morals*', in *Moral Writings*, ed. J. MacAdam (Oxford and New York: Clarendon Press, 2002c), 50–76.
- Stalzer, Kenneth, 'On the Normativity of the Instrumental Principle', Ph.D., Philosophy, Stanford University (2004).

## 11

Resisting the Buck-Passing Account  
of Value

Pekka Väyrynen

## 1. Introduction

T. M. Scanlon's "buck-passing account" of value continues a long tradition of analyzing value in terms of non-evaluative normative notions.<sup>1</sup> Buck-passers about value hold (speaking roughly for now) that to be valuable is nothing more or other than to have other properties that provide reasons for certain positive responses—namely, certain "pro-attitudes" and/or actions expressive of them—to the bearers of those properties. This is to pass the normative "buck" from value onto other properties: the reasons to favor valuable things are provided not by their value but by the properties that make them valuable (Scanlon, 1998: 97). To illustrate, as the prospects of reaching Mordor turn bleak and Frodo Baggins's spirit falters, Samwise Gamgee tries to lift Frodo's mood with an evaluative claim: "There's some good in this world, and it's worth fighting for."<sup>2</sup> According to the format of analysis favored by buck-passers, the fact that something is worth fighting for would just be the fact that it has other properties that provide

I presented an earlier version as "The Buck-Passing Account of Value (Almost Refuted)" at the 2004 Wisconsin Metaethics Workshop and the *Practical Reason and Moral Motivation* meeting in Rome. (Sometimes the direction of progress with work-in-progress is towards wimpier titles.) I am grateful to these audiences for helpful discussion, and to Russ Shafer-Landau for organizing a wonderful workshop. Many thanks to Jonas Olson, Philip Stratton-Lake, Jussi Suikkanen, and an anonymous referee for comments on earlier drafts which led to numerous improvements. I am indebted to Christian Coons for very helpful conversations during the early stages of the paper.

<sup>1</sup> Scanlon introduces the buck-passing account of value in his (1998: 95–8). Other recent proponents of the view include Parfit (2001), Suikkanen (2004), and Stratton-Lake and Hooker (2006). Rabinowicz and Rønnow-Rasmussen (2004) provide a useful overview of the tradition Scanlon continues.

<sup>2</sup> The line is from the movie *The Lord of the Rings: The Two Towers*.