

## CRITICAL STUDY

### SEARCH FOR THE SOURCE

BY MICHAEL SMITH

*The Sources of Normativity*. BY CHRISTINE M. KORSGAARD. (Cambridge UP, 1996.  
Pp. xv + 273. Price £35.00.)

*The Sources of Normativity* is an ambitious and demanding book. The first part comprises Christine Korsgaard's Tanner Lectures, delivered at Clare Hall, Cambridge, in 1992. It constitutes a short but extremely impressive book in its own right. The second part comprises commentaries by Gerry Cohen, Raymond Geuss, Thomas Nagel and Bernard Williams, with a substantial reply by Korsgaard. These papers constitute a high-powered book symposium, the sort of thing that might have appeared in a scholarly journal if the first part had been published separately.

In her lectures, Korsgaard purports to provide nothing less than a grand synthesis of all the supposedly conflicting views that can be taken on the sources of normativity (pp. 164–6). But she has a more partisan goal as well. She tells us that we engage in normative thinking in the practical realm because we have the capacity to reflect on our desires and ask whether we should act in the way we happen to desire. She insists that a certain conception of normative thinking follows from this, and that normative thinking, so construed, when properly conducted, entails a commitment to assigning value to reflective creatures like ourselves in our deliberations. The main line of argument thus has a distinctively Kantian flavour to it: anyone who assigns normative significance to anything at all is supposed thereby to be committed to assigning normative significance to people and so to treating people as ends in themselves.

Korsgaard's lectures are packed with arguments and lessons from the history of philosophy. The conclusions she seeks to establish are, if true, profound, and the

arguments for those conclusions, which she lays out with characteristic flair and ingenuity, are strikingly original. Nor should any of this be surprising. Korsgaard is perhaps the most imaginative moral philosopher writing today. But neither, therefore, should it be surprising to hear that the book makes such significant demands upon its readers. Although Korsgaard attempts to present her grand synthesis in four lectures, there simply is not the space for her to provide all of the details of all of the arguments she needs to give. Readers are thus left with the daunting task of providing those details for themselves. (Excellent though they are, the critical commentaries by Cohen, Geuss, Nagel and Williams did not seem to me to help in this regard. Korsgaard's lectures need and deserve to be read and digested before you read someone else's criticisms. I therefore urge readers to resist reading the commentaries straight away; and the same applies to this review.)

It is impossible to do full justice to *The Sources of Normativity* in a review essay such as this. I shall therefore concentrate on Korsgaard's partisan goal: her defence of a Kantian view about the sources of normativity. It was evidently this part that most excited the commentators when they first heard Korsgaard deliver her Tanner Lectures. I suspect it is the part of the book that will most excite the general reader as well. Certainly it was the part that most intrigued me.

## I. THE STRUCTURE OF KORSGAARD'S MAIN ARGUMENT

Though she would probably dislike its being described in these terms, Korsgaard's argument seems best understood as moving from a meta-ethical premise (the premise about the nature of normative thinking) to a normative conclusion (the conclusion that people have value as ends in themselves). However, the very fact that it has this structure is problematic. It is worth emphasizing the problem, as Korsgaard explicitly attempts to overcome it.

Arguments from meta-ethical premises to normative conclusions are all supposed, by Humeans, to face a dilemma. Either they are unsound, or they turn, at some crucial point, on a stipulation. Someone might argue, for example, from the meta-ethical premise that moral requirements are requirements of rationality to the normative conclusion that we should seek to satisfy the interests of all people affected by our actions. According to Humeans, those who advance such an argument either stipulate a controversial meaning for the word 'rational' – for example, they might define a rational agent to be someone who seeks to satisfy all interests, independently of whose interests they are – in which case the argument will be valid but devoid of relevance (for who says we ordinarily mean *that* by 'rational?'); or else they use the term 'rational' in a more uncontroversial way – for example, they might define rational agents to be those who maximally satisfy their own interests – in which case the argument is unsound (for how could moral requirements be defined in terms of rational requirements in that relatively uncontroversial sense of 'rational?').

Korsgaard's task, as I see it, and therefore the real power of her argument, if it succeeds, is to provide a valid argument from meta-ethical premises to a normative

conclusion without at any point relying on any such stipulation. In order to succeed in this task she must therefore provide a meta-ethical premise that her sceptical Humean opponents would have no antecedent reason to reject. And indeed this is just what she does. At the beginning of her third lecture she makes an observation that is meant to be friendly to both Humeans and Kantians alike (pp. 92–3). As already mentioned, she notes that the fact that we have the capacity to reflect on our desires allows us to ask ourselves whether we should desire as we do. This constitutes what she calls the ‘normative problem’. She then argues that, just as reflection gives rise to the normative problem, so it provides us with the beginnings of a solution. For we answer the question by seeing whether we can reflectively endorse desiring as we do: that is, by seeing whether, on reflection, we can sustain a desire to so desire.

Here, then, we have Korsgaard’s meta-ethical premise. Normative thinking is a matter of the formation and expression of reflective desires about our own state of desire. Moreover, as required, the premise does not stipulate a controversial meaning for ‘normative thinking’. The upshot is thus that if Korsgaard is right that *this* premise entails that we must assign reflective creatures value in our deliberations, then there is indeed an argument from a meta-ethical premise to a normative conclusion, an argument that even her Humean opponents should find it difficult to resist.

## II. THE NATURE OF NORMATIVE THINKING

The view of normative thinking Korsgaard proposes (p. 99) is similar to the second-order desire account of valuing associated most famously with Harry Frankfurt. This account is, however, ambiguous in various ways, and I am not certain which disambiguation she favours.

To begin with, is normative thought supposed to be a matter of the formation and expression of *beliefs* about which second-order desires we would have on reflection? Or is it supposed rather to be a matter of the formation and expression of reflective second-order *desires*? If the former, then Korsgaard is a cognitivist, and the normative question is about a matter of fact. However, she thereby incurs the burden of telling us what these normative facts are. Famously, that is no easy task. If she holds the latter view, by contrast, then she is a non-cognitivist. The normative question is a matter of decision, not a question about a matter of fact. But she thereby incurs the burden of telling us which desires it is that we express when we make this decision. And again, famously, that is no easy task.

Unfortunately Korsgaard does not address this ambiguity. Given that the debate between cognitivists and non-cognitivists has been of the first importance in contemporary meta-ethics, that is too bad. It would have been nice to know which way a Kantian like Korsgaard would like us to go on the difficult analytic questions involved. For the purposes of the discussion that follows I shall therefore simply assume that Korsgaard would prefer the cognitivist reading of Frankfurt’s view. Nothing much of substance should turn on this assumption.

There is another and more pressing problem, however. We are supposing that normative thought is a matter of the formation and expression of beliefs about what we would desire on reflection. But what exactly is meant by 'what we would desire on reflection'? There are at least two possibilities. A desire might be reflective simply in so far as it is an attitude towards one's own state of desiring. On this account, to say that we would desire something on reflection is to say that we would desire to desire it (or perhaps that we would desire to desire to desire it, or ...). Alternatively, a desire might be reflective in so far as it has been formed as the result of a process of reflective critical scrutiny. On this account, to say that we would desire something on reflection is to say that we would desire it if, say, we had a set of desires which was immune to various forms of rational criticism such as being uninformed, or failing to cohere with our other desires, or contributing disunity to an otherwise unified desire set, or....

Moreover, these senses are quite distinct. We might desire something on reflection in one sense without desiring it on reflection in the other. A desire might be second-order and yet disappear if we were to subject our desire set to a process of reflective critical scrutiny; and, conversely, our desires might all be first-order and yet remain if we were to subject them to a process of reflective critical scrutiny. The question therefore naturally arises what Korsgaard means when she talks of what we would desire on reflection.

Again she does not tell us. In this case, however, the choice is far more clear cut. If what we would desire on reflection is to be a source of normativity then it must be read in the 'desires that we would retain if we were to subject our desires to a process of reflective critical scrutiny' way, rather than the 'higher-order desire' way. The reason is familiar from Gary Watson's discussion of Frankfurt's higher-order desire account of valuing. There is no inherent normative feature possessed by desires which are *simply* higher-order. There is, however, an inherent normative feature possessed by desires that survive a process of reflective critical scrutiny, namely, that very feature.

The upshot thus seems to be that Korsgaard is best interpreted as supposing that normative thought is a matter of the formation and expression of beliefs about the desires we would have if our desires were to survive a process of reflective critical scrutiny. Now this might come as somewhat of a surprise. After all, we began with the suggestion that Korsgaard embraces a second-order desire account of valuing similar to Frankfurt's, but the account we have ended with makes no mention of second-order desires. The desires in question might all be first-order. In fact, however, it seems to me that Korsgaard's commitment to a second-order desire account of valuing was always rather superficial, an artefact of a confusion that arises all too easily if we suppose, as she does, that the basic form of the normative question is 'Should I desire that  $p$ '?

After all, when people ask 'Should I desire that  $p$ ?', what question are they asking? The answer is radically unclear. They might be asking whether  $p$  is desirable, or alternatively they might be asking whether desiring that  $p$  is desirable. But these questions are utterly different from each other, different because they can get quite different answers. It might be desirable that  $p$  even though my desiring that  $p$  is not

desirable (a Mad Scientist will kill anyone who desires the things that are desirable), and my desiring that  $p$  might be desirable even though  $p$  is not desirable (a Mad Scientist will destroy the entire universe unless I desire something that is not desirable).

My suspicion is that Korsgaard does not notice any of this. She thinks that the question 'Should I desire that  $p$ ?' is a way of asking only one question, namely, 'Is  $p$  desirable?'. But she then confusedly supposes that we answer that question by seeing whether we would desire to desire that  $p$  on reflection. This is a confusion, because that is in fact the way in which we would answer the rather different question 'Is it desirable that I desire that  $p$ ?'. It therefore seems to me to be best not to invite any of these confusions in the first place. The basic form of the normative question is not 'Should I desire that  $p$ ?', but rather 'Is it desirable that  $p$ ?' for arbitrary ' $p$ ' (including that instance where ' $p$ ' is 'I desire that  $q$ '). And Korsgaard's suggestion, properly understood, must therefore be that we answer this question by seeing whether we would desire that  $p$  on reflection – not whether we would desire to desire that  $p$  on reflection (though if ' $p$ ' is 'I desire that  $q$ ', then we might have to see whether we would desire to desire that  $q$  on reflection).

Does the idea that normative thinking is a matter of the formation and expression of beliefs about the desires we would have on reflection sit happily with the rest of what Korsgaard has to say about normative thinking? It seem to me that it does. Korsgaard sums up her view in the following terms (p. 100):

When you deliberate, it is as if there were something over and above all of your desires, something which is *you*, and which *chooses* which desire to act on. This means that the principle or law by which you determine your actions is one that you regard as being expressive of *yourself*.

But this remains true if normative thinking, and hence deliberation, is a matter of the formation and expression of beliefs about what we would desire that we do on reflection.

Your reflective self, the real you, is the self who has the first-order desires you would have if you were to reflect. When you deliberate, you form beliefs about your real self's desires, beliefs which in turn are able to determine your actions. This is because your beliefs about what you would want yourself to do if you were to reflect have a content that uniquely suits them to cause and rationalize your possession of corresponding first-order desires. It thus follows that, when you deliberate, your reflective self does indeed choose which desires you are to act on, and, when you act on these desires, you do indeed act in ways that are expressive of your real self.

Let us return to the main question, which was whether Korsgaard gives us a way of getting from a meta-ethical premise to a normative conclusion. What I have in effect just done is tidy up her meta-ethical premise. Normative thinking is a matter of the formation and expression of beliefs about what we would first-order desire if our desires were to survive a process of reflective critical scrutiny. The question we must address is whether it follows from this meta-ethical premise that people should always be treated also as ends, never merely as means.

## III. PRACTICAL IDENTITIES

Korsgaard's argument for this conclusion is subtle.

She begins by arguing that there is a tight connection between the answers we give to the questions that reflection poses for us in the practical realm and the conceptions we have of our practical identities. For example, I might wonder whether I should look after my children in my current circumstances. That is, I ask myself whether looking after my children in my current circumstances is something that, on reflection, I would desire myself to do in these circumstances. In order to answer this question I must form a belief about the substance of my reflective self's cares and concerns. The content of this belief is my conception of my practical identity. If I believe that my reflective self would want nothing more than for me to look after my children in my current circumstances, then it follows that I conceive of my reflective self as, to that extent, a devoted parent. This description of my reflective self conveys the substance of his cares and concerns.

This, in turn, is important, because Korsgaard suggests (p. 101) that the conception I have of my practical identity is a description of my life under which I value myself:

The conception of one's identity in question here is not a theoretical one, a view about what as a matter of inescapable scientific fact you are. It is better understood as a description under which you value yourself, a description under which you find your life to be worth living and your actions to be worth undertaking.

Thus if I conceive of my reflective self as, in the sense described, a devoted parent, according to Korsgaard it follows that I think that being a devoted parent, in the sense described, is what makes my life worth living and my actions worth undertaking.

Given the importance that attaches to this claim in her overall project, we might have hoped that Korsgaard would provide an argument for it. Unfortunately, however, as far as I can see she provides no argument whatsoever. She simply asserts that our practical identities provide us with descriptions of our lives under which we value leading them. We must therefore ask what might be said in support of the claim.

There are two ways of understanding talk of reflective selves who stand over and above us when we act, urging us to act in one way rather than another. On the one hand we can conceive of our reflective selves as providing us with examples of behaviour that we should try to emulate. On the other hand we can conceive of them as providing us with advice that we should try to follow. In her reply to Williams on internal and external reasons, Korsgaard seems to commit herself to the view that we should interpret such talk on the model of exemplars, rather than as advisers. The conception of our reflective selves as exemplars is exactly the premise that Korsgaard needs at this point in her argument.

If, when I judge it desirable to look after my children in my current circumstances, I thereby express not just my belief that my reflective self would want me to

act in this way in these circumstances, but my belief that my reflective self would himself act in this way in these circumstances, and if I have to suppose that my reflective self thereby provides me with an example of behaviour that I should try to emulate, then it does indeed seem to follow that I suppose that being like my reflective self is what would make my life worth living and my actions worth undertaking. An exemplar is, after all, someone I admire, someone I aspire to be like.

I shall return to the distinction between thinking of our reflective selves as exemplars and thinking of them as advisers presently. For now, I simply note that if we have to think of our reflective selves as exemplars, then Korsgaard has the support she needs for her claim that our practical identities provide us with descriptions of our lives under which we value leading them.

She notes that the mere fact that I have a certain conception of my practical identity does not mean that I should conceive of myself in that way. And that is of course true. The mere fact that I believe that my reflective self has the cares and concerns of a devoted parent does not mean that this belief is true. If it is false – if, say, my reflective self has the cares and concerns of a devoted parent only up to a point, say up to the point where providing for my children would require other people's children to bear significant burdens – then I should not conceive of my reflective self as a devoted parent *simpliciter*. I should rather conceive of my reflective self as a parent who devotes himself to his children provided that doing so does not make other people's children bear significant burdens. The crucial question (p. 103) is therefore 'How should I conceive of my practical identity?'

At this point Korsgaard makes a crucial move in her argument. She notes that we shall have an answer to this question if we can find a conception of our practical identity that we cannot legitimately question. Now, as we have already seen, we can most certainly question many of our practical identities. For example, I can question whether I really have the practical identity of a devoted parent, and it seems that I could equally question whether I have the practical identity of a lover, or a friend, or an academic, or a citizen of Australia, or .... The mere fact that I can and do conceive of myself in each of these ways – that is, the fact that I believe that I would want myself to act as a lover, a friend, an academic, a citizen of Australia, ... , if I had the set of desires that survived a process of reflective critical scrutiny – does not entail that I should. But if I could find a conception of the cares and concerns of my reflective self that it made no sense for me to question, then the mere fact that I could conceive of myself in this way would entail that I should.

Korsgaard's interesting suggestion, at this point, is that there is in fact one such conception of my practical identity, for I cannot question whether or not I have the practical identity of a creature who is capable of reflective questioning of his desires (pp. 103–25). That is to say, when I look for a description of the cares and concerns of my reflective self, I cannot doubt that it would be appropriate to describe my reflective self as having the cares and concerns of a creature who is capable of forming desires as the result of reflection. It is, after all, in the nature of my reflective self *as a reflective self* to have such cares and concerns. It therefore follows that I cannot question this conception of my practical identity. It makes no sense at all to

ask whether I should have this practical identity, because I cannot have any practical identity at all without having this one.

It should now be plain why Korsgaard thinks that her meta-ethical premise entails a normative conclusion. For once we add this suggestion to her earlier claim that my practical identity provides a description of my life under which I find it worth living and my actions worth undertaking, we can draw the strictly normative conclusion that if I value anything at all then I must value being a creature who has desires as the result of the exercise of the capacity to form desires via reflection. And since this argument has not turned on anything special about me and my concerns, as opposed to those that would be had by other people, we can suppose that if it is valid at all then it is valid for other people too. If each of us is to value anything at all then we must each value being a creature who has desires as the result of the exercise of the capacity to form desires via reflection.

#### IV. OTHER PEOPLE

Korsgaard recognizes that this conclusion, though normative in nature, falls short of Kant's conclusion that people in general must always be treated as ends in themselves, never merely as means. If I must value being a creature who has desires as the result of the exercise of the capacity to form desires via reflection, then perhaps it follows that I must treat myself always as an end, never merely as a means. But it is consistent with my doing so that I treat other people merely as means to my own ends. I may subordinate them in my attempt to realize my own perfection. Likewise, all other persons may subordinate me along with the rest in their attempts to realize their own perfection. Another argument is therefore needed to get from this egoistic, albeit normative, conclusion to Kant's more altruistic conclusion.

At this point (p. 143) Korsgaard adapts an argument of Nagel's from *The Possibility of Altruism*:

How does the obligation come about? Just the way Nagel says that it does. I invite you to consider how you would like it if someone did that to you. You realize that you would not merely dislike it, you would resent it. You would think that the other has a reason to stop, more, that he has an obligation to stop. And that obligation would spring from your own objection to what he does to you. You make yourself an end for others; you make yourself a law to them. But if you are a law to others in so far as you are just human, just *someone*, then the humanity of others is also a law to you. *By making you think these thoughts*, I force you to acknowledge the value of *my* humanity, and I obligate you to act in a way that respects it.

Thus, according to Korsgaard, Nagel's argument shows not just that *you* should value *my* being a creature who has desires as the result of the exercise of the capacity to form desires via reflection, but also that *I* should value *your* being a creature who has desires as the result of the exercise of the capacity to form desires via reflection.

If this argument is successful, then it follows that we can strengthen the earlier egoistic conclusion. The new conclusion is that if each of us is to value anything at

all, then we must each value creatures who have desires as a result of the exercise of the capacity to form desires via reflection. The conclusion is therefore just the one advertised at the outset. We have moved from premises about the nature of normative thinking as such to the Kantian conclusion that if anything is valuable at all, then creatures who have desires as the result of the exercise of the capacity to form desires via reflection are valuable in themselves. Our capacity to form reflective desires is the source of normativity in both the meta-ethical sense (for we can analyse valuing something in terms of believing that we would desire that thing if we were to reflect) and in the normative ethical sense (for creatures with the capacity to form such reflective desires have value). So, at any rate, Korsgaard would have us believe.

## V. EVALUATION OF KORSGAARD'S MAIN ARGUMENT

I said earlier that there is an undefended premise in Korsgaard's main argument. It is time to subject that premise to closer scrutiny.

Korsgaard claims that our conceptions of our practical identities provide descriptions of our lives under which we find them worth leading and our actions worth undertaking. I noted that this premise requires some sort of defence, and that Korsgaard could provide the needed defence if she were to appeal to a view she seems committed to elsewhere, namely, the view that when we deliberate, and imagine our reflective selves over and above us choosing which desires we are to act upon, the imagined choices of our reflective selves provide an example we are to emulate. If our reflective selves provide an example for us to emulate, if their choices are ones we have to suppose ourselves aspiring to make, then we must indeed value leading lives like those our reflective selves lead, and the choices that they make.

It will come as no surprise to hear that this is where the crucial move from meta-ethical premise to normative conclusion takes place. In fairness, it must be said that it does not look like any sort of stipulation. The move seems licensed not by a stipulation, but instead by our apparent inability to remain steadfastly neutral on all normative ethical issues when we attempt to do meta-ethics. If the judgement that it is desirable to act in a certain way is an expression of your belief that your reflective self would want you to act in that way, and if, as the example interpretation of such talk requires, we have no alternative but to interpret this as the belief that your reflective self provides an example for you to emulate, then the only conclusion to draw is that the attempt to do value-free meta-ethics leads you ineluctably back into doing value-laden normative ethics. The meta-ethical argument itself forces the normative assumption upon you. You do not stipulate anything.

The situation thus seems to be that if Korsgaard is right, that if when we deliberate, and imagine our reflective selves over and above us choosing which desires we are to act upon, we have to suppose that the imagined choices of our reflective selves provide us with an example we are to emulate, then her argument might well go through. But though this might all sound quite plausible initially, I think that in the end we must reject it. We must reject it because the idea that our reflective selves

are people whose behaviour we are to emulate is simply implausible. We must suppose instead that our reflective selves are people whose advice we are to follow.

Why must we reject the idea that our reflective selves are exemplars whose behaviour we are to emulate? Ironically, the main reason is provided by Korsgaard herself. If that idea were plausible then, much as Korsgaard says, valuing anything at all would commit me to valuing having cares and concerns like those that I believe my reflective self has. But now let us translate that claim using Korsgaard's own analysis of what it is to value something. This is what we get: believing that my reflective self would want me to act in a certain way in certain circumstances commits me to believing that my reflective self would want that I have cares and concerns like those that he himself has. But that does not seem in the least plausible.

The only thing we know for certain about my reflective self's cares and concerns is that they are one and all reflectively formed. But in that case, since it is conceivable that my reflective self should want me – indeed, since it is conceivable that he should want himself – to have all sorts of cares and concerns spontaneously, or naturally, or at any rate not as the result of their being reflectively formed, it follows that it is conceivable that my reflective self should want me – indeed, it is conceivable that he should want himself – to have cares and concerns very different from those which he has.

The desires of the devoted parent again provide an example. It is surely conceivable that my reflective self should want me – indeed, it is conceivable that he should want himself – to have care and concern for my children as a natural response to their needs, rather than on the basis of reflection, the way in which he has them. Perhaps children can detect the difference between a concern for them that is naturally formed in response to their needs, as against a concern for them that is reflectively formed. Perhaps they develop in a happier and a healthier way when those who care for them have the former cares and concerns, rather than the latter. But if this is right, then my reflective self certainly would not want me – nor, indeed, would he want himself – to have the desires that he has. For his desire that I look after my children is formed not as a natural response to the needs of my children, but rather on the basis of reflection.

Plainly we could generate more counter-examples. The general idea is simply that it is conceivable that my reflective self should want me – indeed, it is conceivable that he should want himself – not to be reflective. Perhaps people who are reflective lead miserable lives, as opposed to those who are unreflective, or, worse still, perhaps the very fact that they are reflective makes everyone lead a miserable life. What would my reflective self want me to do in circumstances in which a Mad Scientist will blow up the world if I acquire the desires that my reflective self has? My reflective self would then presumably want very much for me to be unreflective. It is easy to imagine his horrified reaction if I followed his example, and, as the result of reflection, acquired the desires that he has (including the desire that I be unreflective!). As I embarked on this process, he would no doubt be hoping that I would stop and do just what he wants me to do, that is, not reflect at all. In such cases my reflective self certainly would not want me – nor, indeed, would he want himself – to lead a life like the one he leads.

It might be thought that Korsgaard could avoid this conclusion. She might argue that while I am not to follow the example of my reflective self, I am required to follow the example I set for myself in an idealized possible world – specifically, in the world in which I am doing exactly what my reflective self wants me to do. In this sense, she might say, I do have to treat the behaviour of my idealized self – the one who follows my reflective self’s advice, if not my reflective self himself – as an example that I am to follow.

But though it is true that she could in this way turn the view that our reflective selves offer us advice into a view according to which the behaviour of one of our idealized selves provides us with an example, the cost of doing so is to lose the crucial feature of the *pure* example model that allowed Korsgaard to derive her substantive normative conclusion, the conclusion that I have to value my own reflective nature. The crucial feature was not just that I have to aspire to be like my exemplar, but that, when my exemplar is my reflective self, I have to aspire to be like someone whose cares and concerns are the product of his capacity to form desires via reflection. But when my exemplar is an idealized self who is simply doing what my reflective self wants him to do, then, given that my reflective self might want him not to be reflective, he might not be being reflective. Korsgaard would then no longer be possessed of a premise from which she could draw the conclusion that I have to value my own reflective nature.

The objection does not depend on interpreting her as a cognitivist. Suppose it is a set of my second-order desires, those expressive of my identity as a creature capable of forming desires on the basis of rational reflection, that are immune from rational criticism. What the counter-examples just given show is that when I exercise this capacity I would not necessarily desire to have and express this capacity.

What all this shows, I think, is that it is simply not true that the non-optional conception of my practical identity that I share with everyone else, my conception of myself as a creature who is capable of forming desires via reflection, provides a description of my life under which I value it. Maybe it does; maybe it does not. It all depends on the substance of my reflective self’s cares and concerns. Since my reflective self might not want me, or even himself, to have the reflectively formed cares and concerns that he has, it is simply wrong to suppose, as Korsgaard does, that I must value leading a life like the life that my reflective self leads. My reflective self need not provide me with an example that I am to emulate. The crucial premise that enables Korsgaard to move from meta-ethical premise to normative conclusion is therefore one that we have to reject.<sup>1</sup>

*Australian National University*

<sup>1</sup> Earlier versions of this paper were read at a memorial conference in honour of Warren Quinn at UCLA, to a meeting of the Jowett Society at Balliol College, Oxford, and at the Department of Philosophy Staff Seminar, Sheffield University. I would like to thank everyone who participated in discussions on those occasions. I am especially grateful to David Archard, Jules Coleman, Tim Hall, Barbara Herman, Brad Hooker, Vera Koffman, Philip Pettit, Charles Pigden, Mike Ridge, Geoffrey Sayre-McCord, Mike Thau, Valerie Tiberius and R. Jay Wallace.