

Moral Character and the Iteration Problem

GARRETT CULLITY
University of St. Andrews

Moral evaluation is concerned with the attribution (to its various objects – actions, character, attitudes, states of affairs, institutions) of values whose distinction into two broad groups has become familiar. On the one hand, there are the most general moral values of rightness, wrongness, goodness, badness, and what ought to be or to be done. On the other, there is a great diversity of more specific moral values which these objects can have: of being a theft, for instance, or a thief; of honesty, reliability or callousness.¹ Within the recent body of work attempting to restore to the virtues a central place in ethical thinking, two claims stand out. One is that, of these two kinds of values, the specific ones are explanatorily prior to the general – that if an action is wrong, it is because it is wrong in one of those specific respects.² A second claim, though, is now standardly made definitive of ‘virtue ethics’: that amongst the specific values, the value of character is explanatorily prior to that of action – that if an action is callous, say, it is because it expresses callousness of character – and that in this sense, the moral value of action derives from that of character.³ This second claim has been widely attacked; in what follows, I present a reason for believing that, at least in the case of callousness, it is right.

In doing so, I shall be using ‘callousness’ with a wider reference than some readers will want to give the term. For I shall not be restricting it to cases where an agent is strictly unfeeling – where he simply fails to engage sympathetically with the predicament of another person.

¹ For Bernard Williams’s distinction between ‘thin’ and ‘thick’ ethical concepts, see *Ethics and the Limits of Philosophy*, London, 1985, pp. 129–30, 140–1.

² See e.g. Iris Murdoch, *The Sovereignty of Good*, London, 1970, p. 43; John McDowell ‘Virtue and Reason’, *The Monist*, lxiii (1979), 331; S. L. Hurley, ‘Objectivity and Disagreement’, *Morality and Objectivity*, ed. Ted Honderich, London, 1985, pp. 54–97; Garrett Cullity, ‘Aretaic Cognitivism’, *American Philosophical Quarterly*, forthcoming.

³ See e.g. Kurt Baier, ‘Radical Virtue Ethics’, *Midwest Studies in Philosophy*, xiii (1988), 127; David Solomon, ‘Internal Objections to Virtue Ethics’, *Midwest Studies in Philosophy*, xiii (1988), 428–9; Sarah Conly, ‘Flourishing and the Failure of the Ethics of Virtue’, *Midwest Studies in Philosophy*, xiii (1988), 84–5; and Phillip Montague, ‘Virtue Ethics: A Qualified Success Story’, *American Philosophical Quarterly*, xxix (1992), 53. For both claims, see William K. Frankena, *Ethics*, Englewood Cliffs, 1973, p. 63. A third ‘anti-Kantian’ claim sometimes identified as definitive of an ‘ethics of virtue’ is the denial ‘that it is a necessary condition of perfectly moral personhood that one be governed by a concept of duty’. See Marcia Baron, ‘Varieties of Ethics of Virtue’, *American Philosophical Quarterly*, xxii (1985), 47.

Clearly, callousness does centrally include such cases; but I shall use the term, more broadly, to include those cases where an agent possesses the sympathetic feelings, but fails to act on them – thanks to an overriding concern with his own interests, perhaps, or obedience to authority – and is to be morally criticized for so failing. According to my broader usage, then, callousness is a morally objectionable failure to have motivationally efficacious sympathetic feelings. Anyone who insists on the more restricted usage should simply substitute a different evaluative term to cover the range of moral failings which I am covering with the term ‘callousness’.

I

Before presenting my reason for the virtue-ethical claim concerning callousness, it is worth considering the obstacles in the way of the idea that there can ever be this relation of priority of the moral value of character over that of action. I shall concentrate on two of the most serious challenges.⁴

The first is that it is plainly wrong to infer that an action is not callous from the fact that the agent performing it is not a callous person (or vice versa). Actions can be performed ‘out of character’: someone who is not a callous person can perform callous actions on occasion. Stanley Milgram’s famous obedience experiments, in which subjects were induced by an authoritative instructor to do what they believed to amount to administering extremely painful and even lethal electric shocks to their ‘victims’ (in fact, actors), seem clearly to illustrate this.⁵ They illustrate not the surprising prevalence of callous character, but the surprising extent to which uncallous but obedient people can be induced to perform callous actions. Conversely, however, one can infer from the fact that someone regularly performs callous actions when the opportunity presents itself that he is a callous person. If anything, it looks as though the moral value of character derives from that of action.⁶

Clearly, this point applies quite globally: for any feature *F* of character, I can act *F*-ly on occasion without being an *F* person. How should the virtue-ethicist reply? The best strategy seems to be to exploit the following thought: although an occasional impatient action (say) may

⁴ For other serious challenges, see Gregory W. Trianosky, ‘Supererogation, Wrongdoing, and Vice: On the Autonomy of the Ethics of Virtue’, *The Journal of Philosophy*, lxxxiii (1986), 26–40; and Conly, ‘Flourishing and the Failure of the Ethics of Virtue’, 83–96.

⁵ Stanley Milgram, *Obedience to Authority: An Experimental View*, New York, 1974.

⁶ It is no good trying to rescue the explanatory claim by saying that a callous action is one that *would* be performed by a callous person, either. Such a person would only do so if he were acting in character.

not issue from an impatient character, there will always be a sense in which it expresses an impatient element in the agent's character. My performing an impatient action shows that part of my character – the part expressed in this action – contains the ingredients distinctive of impatience; what would be required to make it appropriate to call me an impatient person is simply that this is a significant enough part of my character to be mentioned in an overall summary. On this suggestion, a proponent of the virtue-ethical claim concerning callousness must produce an explanation of how an action can *express* callousness of character without the agent's being a callous person. But obviously, such an explanation will only meet the objection if the criterion for an action's expressing callousness of character is not simply that the action itself is callous.

To approach the second objection, consider the following distinction between the moral values that actions can have. Some can be possessed by an action irrespective of the attitudes of the agent towards it: an action can sometimes be unfair, or honest, it seems, independently of the agent's reasons and motives for, intentions in, and beliefs about, its performance.⁷ (Let us call mental states such as these, whose intentional object is an action, *action-directed attitudes*.) But others, it seems, cannot: an action can only be kind, courageous or vain if the agent has the appropriate attitudes towards it. This gives us a distinction amongst the values of action between the *attitude-independent* and the *attitude-dependent*.

Now I cannot see any sensible ground for thinking that attitude-independent values of action ever derive from values of character. For here, it is not simply that dishonest people sometimes perform honest actions – an honest action need not even *express* honesty of character in any way. It seems that honesty of character must be explained as the disposition to perform honest actions from the right attitudes. So the strongest virtue-ethical claim with any plausibility is that all attitude-dependent moral values of action – values, that is to say, like callousness – derive from those of character. How might one support this claim? The following route suggests itself.

A simple but adequate account of what constitutes an agent's character would appear to be this: it is the total, relatively settled pattern formed by his various attitudes (especially the action-directed ones), his potential and actual actions, and the relation between them and their surrounding circumstances.⁸ In making judgements about a

⁷ Compare Kant, 'Groundwork of the Metaphysic of Morals', *Kant's gesammelte Schriften*, vol. iv, Berlin, 1903, pp. 397–8.

⁸ It seems to me that an unexercised disposition is not part of a person's character. Anyone who disagrees with this, though, can replace the suggested account of character with one referring to the dispositions which give rise to the pattern I mention without affecting any of the subsequent arguments.

person's character, it is clear that we are often generalizing about the sorts of attitudes he forms and actions he performs in different situations. But potential actions which are never performed can also be relevant, in the relation they bear to an agent's attitudes, to the assessment of his character – most obviously, to whether it includes strength or weakness of will. It is also easy to see why I need to talk about a 'total, relatively settled pattern': this accommodates the possibility of acting out of character which has already been mentioned. If we take a snapshot of action, attitudes, and circumstances at a particular moment, we may find that the relations between them are distinctive of impatience (say), even if the agent is not an impatient person. Of course, if one typically becomes impatient in a specifiable range of situations, then this will be a feature of character, even if impatience *tout court* is not; but my simple account of character accommodates this point, since the former feature does plausibly belong to the total, relatively settled pattern it mentions. Moreover, if the impatient episodes are simply unpredictable or erratic, then this unpredictability is surely also naturally thought of as a feature of character, and is accommodated by the account in the same way. However, although it is only the settled features of a person that we count as features of his character, it must not be forgotten that a person's character can change – and sometimes very suddenly, as in cases of religious conversion.⁹ So it is only *relatively* settled features that are required.

But if this simple account of character – as consisting in attitudes, actions and the relations between them and circumstances – is right, then this suggests a very straightforward way of supporting the virtue-ethical claim concerning attitude-dependent values of action. To say of my action that its value is attitude-dependent is to say that its value depends on the relation between it, the attitude it expresses and the surrounding circumstances. But if so, it depends upon that which, if typical of me, would be properly cited in an overall description of my character. Moreover, even if the expression of this attitude in such an action is not typical of me – even if I am acting out of character – it is still a constituent element in the complete pattern of attitudes, actions, and the relations between them and circumstances which comprises my character, and the action does express this element. And this gives us the virtue-ethical conclusion: to say that the moral value of an action is attitude-dependent is to say that it depends on the value of elements of character which the action expresses.

However, this only serves to introduce the second objection. An opponent of virtue ethics need not deny that in the case of values like

⁹ See e.g. William Sargant's pioneering study of such cases, *Battle for the Mind: A Physiology of Conversion and Brain-Washing*, London, 1957.

callousness, the value of actions depends or supervenes on that of (elements of) character, in a standard sense: two actions can only ever differ in the former respect if they differ in the latter. But this can be conceded while denying that the considerations just given in its support will sustain any conclusion concerning explanatory *priority*. Grant that some values of action are attitude-dependent – that is, that moral significance attaches to certain relations between actions and action-directed attitudes. And grant that facts about a person's character are simply generalizations about such relations, amongst others. This does not yet commit us to a single correct order of evaluative explanation: on the contrary, it seems to show that there is none. We can begin by describing the pattern of relations between action-directed attitudes and actions which is distinctive of a virtue or vice of character, and then say that an action is virtuous or vicious in the same respect if it expresses such a character; or we can begin by explaining the virtuousness or viciousness of an *action* in terms of its relation to the appropriate action-directed attitudes, and then explain the virtue or vice of character as the settled tendency to perform such actions. The relation of dependence between action- and character-evaluation is mutual.

This gives us two simple but powerful challenges to the character-based view of callousness. Now let us turn to my reason for endorsing it none the less.

II

To present that reason, I begin by describing the Iteration Problem.

If an agent cannot be bothered to take a few minutes to save the life of someone dying in front of him, and there are no extenuating circumstances, then I take it we have a paradigm of callousness. This judgement need not be accorded any special privilege as unchallengeable – difficult though it is to imagine what could be a reason for abandoning it – but we certainly begin any discussion of callousness from here. If the life-saving action would involve a serious injury, threat, or loss to the agent, then most of us would tend to regard it as heroic, and the failure to perform it as understandable;¹⁰ but if not, and if there are no other extenuating features of the situation, that failure is surely callous. (Of course, only to say this leaves open the question *how* serious a sacrifice must be to ensure that refraining from saving a life at that cost is not callous. But let us leave this open for the time being: its importance comes later.)

¹⁰ For some dissenting voices, though, see Peter Singer, 'Famine, Affluence and Morality', *Philosophy & Public Affairs*, i (1972), 241, and Shelly Kagan, *The Limits of Morality*, Oxford, 1989.

Now consider a different situation: one in which the agent is faced by an enormous number of dying people, each of whom could be saved at a few minutes' expense. One way of thinking about this case is as the iteration of the first one many times over, so that what is morally demanded of the agent is the conjunction of what would have been morally demanded of him had he been confronted by each of the dying people singly.¹¹ And if we characterize the case in this iterative way, it leaves us with the following, striking conclusion. Not only would it be callous for him to save no-one's life, it would be callous to stop saving lives before the cost of saving one further life became great. No doubt, if he spent a large enough proportion of his time doing nothing but saving lives, the opportunity to spend even a few minutes on something else would start to assume substantial importance. But the upshot of the iterative approach is that doing less than this would be callous.¹² Of course, it is natural to complain that, given someone who saves many lives but then stops doing so before reaching this extreme, we would not say that he simply cannot be *bothered* saving the next person, nor that there are no extenuating circumstances – so the case for holding that he meets the paradigm of callousness falls down. But a proponent of the iterative approach will attack this. Presumably, what deters us from saying that he cannot be bothered saving the next person is that he can plausibly say that he would like to do so, but needs to live his own life as well. However, according to the iterative approach neither this nor anything else will count as a justification for letting this person die unless it would justify abandoning a lone dying person. And it seems clear that someone who said this when presented with a single life to save – that he would like to do so, but must live his own life – would still be paradigmatically callous.

If there is an answer to this challenge, it must lie in advocating an alternative, 'aggregative' view of the assessment of callousness in such a situation. Although it would be callous for an agent to help none of the many dying people, it would not be callous if he responded by assessing, first, the scale of the overall need being presented to him, and secondly, the sacrifice he could reasonably be asked to make in relation to that collective predicament. Once his aggregate sacrifice in saving successive lives reached a certain magnitude, refraining from saving a further life would no longer be callous, even if the sacrifice involved in saving one more person remained small. (Just what is that magnitude? Again, let us note the importance of this question as we

¹¹ An action is morally demanded of an agent, on my usage of this phrase, whenever it would be morally wrong for him not to perform it.

¹² This iterative approach seems to be assumed in Peter Singer's argument for the conclusion that 'we ought, morally, to be working full time to relieve great suffering of the sort that occurs as a result of famine or other disasters' (Singer, 238).

pass, and return to it.) If one takes this aggregative view, the callousness of an individual action can depend on its relation to an agent's policy concerning a range of actions to which it belongs. If I have saved many other lives already, then it may not be callous to refrain from saving one more.

It seems clear that if anything, we tend intuitively towards the latter, aggregative view of the assessment of callousness in such situations, as opposed to the iterative one first described. This does not conclusively establish what we *ought* to adopt this view, but on most plausible moral epistemologies, it is at least a piece of evidence for it. This is the starting point of my argument.

Recognizing this, however, may not seem to have advanced me towards the conclusion that callousness of action derives from that of character. Where the intuitively supported aggregative view differs from the iterative one is in maintaining that the callousness of one action of letting someone die can depend on whether one has already performed other life-saving actions. But the fact that the moral value of an action can depend upon other actions already performed by the agent can hardly be a point of contention for opponents of character-based assessment of action: everyone agrees, for instance, that the wrongness of my performing an action can depend on whether I have promised not to perform it. The moral value of an action is certainly dependent on the context in which it is performed, including its coming after other actions of mine; but why should it be thought to depend on character?

III

The first point to be made in reply is this: it is not as if the callousness or non-callousness of my current action will ever depend simply on whether there are other actions of mine, belonging to appropriate kinds, which precede it. In a situation of the kind being imagined, in which many lives could be saved, an agent may come to form beliefs about the magnitude of the sacrifice which he ought to make on behalf of the dying people, and to believe that it would not be unreasonable to stop saving lives after having made a sacrifice of that magnitude. According to the aggregative view of callousness, if the magnitude in question is large enough, then acting on the resulting attitude will not be callous. But one must really be acting *on* that attitude – the action must be *explained* by it – rather than merely in a manner that happens to conform with it. To see the need for the explanatory relation, consider a racist who, having made a sacrifice of such a magnitude in saving a certain number of lives, refrains from saving the next person because he is black. In acting on the attitude that black people are not

worth saving, and not on the attitude that he has sacrificed enough in saving lives to make it reasonable to stop doing so, the racist surely still acts callously – even if he has the latter attitude, and the attitude itself is reasonable. For an action not to be callous, it is not enough that it should conform with a non-callous policy: that policy must be the agent's policy, and its being the agent's policy must help to explain the action.¹³

By itself, this first point would be simply reminding us of the contrast between callousness as an attitude-dependent value of action and attitude-independent values of action like honesty and justice. But now let us add a second. In situations of the kind being imagined, the attitude on whose value the callousness of an action can depend, if the aggregative approach is right, is an attitude of preparedness to perform those life-saving actions which jointly involve a certain maximum sacrifice. That is, the evaluation of an action as callous or non-callous is certainly attitude-dependent, but it can depend not upon the agent's attitude towards that particular action, but upon his attitude towards a range of actions to which it belongs.

But clearly, an action's callousness will not depend simply on its explanatory relation to some attitude of this kind. It will depend also on the details of the particular attitude in question – specifically, on the magnitude which the agent is prepared to accord to the maximum sacrifice in question. For surely such attitudes can themselves be judged in terms of callousness. Someone who is willing only to make sacrifices with no monetary cost, for instance, would seem clearly to have a callous attitude concerning this maximum sacrifice, in making it too small; conversely, whether you think it is morally admirable or not, a willingness to give up one's life for a stranger certainly is not *callous*. If I were claiming to offer a full account of callousness here, I would need to answer the question I have been postponing until now: What is the magnitude, in between these extremes, that an agent must attribute to the maximum sacrifice he is required to make in response to a given collective need in order to have a non-callous attitude in relation to that need? However, for my present purposes, what is important is simply the *existence* of this question, and not the provision of my own answer to it. All I need to have established is that answering this question – a question concerning the assessment in terms of callousness of attitudes towards ranges of action – is prior to the assessment of individual actions in those terms. Attitudes of this

¹³ It must also explain it in the right, non-deviant way. If, having acted callously, I am remorseful, then presumably a full explanation of any remorseful action of mine will include the explanation of the callous action which gave rise to it, but the remorseful action is not callous. I shall not try to spell out here, though, what non-deviancy consists in.

kind can be evaluated as either callous or not, and they must be so evaluated before we can settle the question of the callousness of the actions explained by them.

Evaluation of these attitudes towards ranges of actions is not yet the evaluation of any individual action within that range. Having morally assessed an attitude of this kind, we can only go on derivatively to assess an action within the range to which it applies if we also know which other actions within the range have been performed. This contrasts sharply with the assessment of attitudes which are directed towards individual actions. There, any potential or actual action non-deviantly explained by such an attitude must be accorded the same moral value as the attitude. And this is not so much a matter of an entailment of action- by attitude-assessment: it is rather that there is really only one object of assessment involved. Whenever an action is explained by the agent's attitude towards it, the attitude can simply be brought within the scope of a broader description of the action. The assessment of the attitude is simply the assessment of an action described broadly enough to include it.

So the evaluation of attitudes towards ranges of actions is not yet the evaluation of individual actions, but is prior to it. Notice also, however, that my claim does not reduce to the more common one that the value of actions derives from that of their motives.¹⁴ An agent's *motive* for ceasing to save lives when he does, and not saving the next person, need only be that he wants to do other things instead. The full explanation of his acting on this motive, though, may include his believing that the aggregate sacrifice he has made for the others makes this permissible.¹⁵

However, this prior evaluation of attitudes, although it is neither the evaluation of actions nor of motives, certainly is the evaluation of some of the constituents of the complete pattern formed by an agent's attitudes, his actions, and the relations between them and the surrounding circumstances. It is the evaluation, in short, of elements of agents' character. Therefore, in this sort of case, the value of character is explanatorily prior to that of action.

IV

How does this answer the two objections to virtue ethics in Section I?

The first was to explain how its claim concerning the priority of character- to action-assessment can be made consistent with the fact

¹⁴ See e.g. Hume, *A Treatise of Human Nature*, ed. L. A. Selby-Bigge, 2nd edn., Oxford, 1978, II.ii.1.

¹⁵ Compare: my *motive* for shooting at an intruder is unlikely to be my thinking that injuring people in self-defence is morally permissible.

that callous actions need not be performed by callous agents. We could do so, I observed, if we could explain what the expression of callousness of character could consist in, without relying on a more basic notion of callous action. In what followed, I have been pursuing the straightforward way of doing so suggested in Section I: callous actions can be callous because they express callous attitudes, which are constituent elements of the agent's character. And this does give us what we ought to say about the subjects of Milgram's experiment and, by extension, about all cases of acting 'out of character'. If I am so constituted as to act under authoritative direction with the intention of causing severe pain to helpless and unwilling victims, *this* part of my character includes the pattern of relations between action-directed attitudes and actions which is distinctive of callousness. If I act in this way, I am expressing an attitude which would, if typical enough of me to feature in a description of my total, relatively settled pattern of attitude-formation, secure my callousness of character.

Now it is natural to want to protest against this that it is a mistake to describe the Milgram-subject as even partly callous. What the experimental result reveals is that he has a disposition to do what he is authoritatively told to do, callous actions included along with all the other possible instructions he might receive: it is not that he is callous, to any extent, but that one is obedient, to a very great extent. But the mistake is to see this as an objection. All that is required for something to count as an element of an agent's character is that it is a constituent of the total, relatively settled pattern formed by his attitudes, his actual and potential actions, and the relations between them and their surrounding circumstances. And if so, there is no obstacle to counting as elements of character those morally relevant attitudes formed by an agent which influence his actions on occasion, even if they are not properly cited as even partial answers to the request for a description of his character. There can be elements of my character whose expression in a particular action makes it callous, even though it is not appropriate to say of me generally that I am (even partly) a callous person.

The second objection was that if an agent's character simply consists in this pattern of attitudes, actions, and relations between them and circumstances, then for actions whose value is attitude-dependent, there is no reason to say that either the value of action or that of character is explanatorily prior. What we have found is that that objection should only seem convincing to someone who has made an oversight – the oversight of neglecting to notice that the action-directed attitudes which contribute to one's character need not all take individual actions as their intentional objects. We have found some intuitive evidence that attitudes whose objects are *ranges* of actions are relevant to

callousness – specifically, attitudes of willingness to make a certain maximum sacrifice in response to a given collective need. The crux of my argument has been that sometimes we need already to have evaluated such attitudes as callous or not before evaluating individual actions in those terms. To evaluate attitudes of this kind as callous is not yet to evaluate action; but it is to evaluate an element of the agent's character. Therefore, the second objection fails too. There is at least one kind of case where the virtue-ethical claim concerning the priority of character- over action-evaluation must be accepted: a kind of case, and a particular vice, for which we cannot first assess individual actions, and then say that the vice of character is the disposition to perform such actions.¹⁶

¹⁶ This paper has been much improved by the helpful comments of Berys Gaut, John Haldane, and Brad Hooker.