

## ARE WE RESPONSIBLE FOR OUR CHARACTERS?

Neil Levy (University of Melbourne)

nlevy@unimelb.edu.au

**Abstract:** A number of philosophers have argued in recent years that we are each, typically, responsible for our characters; for what we are, as well as what we do. This paper demonstrates that this is true only of the basically virtuous person; the basically vicious are not responsible for their characters. I establish this claim through a detailed examination of the conditions upon the attribution of moral responsibility. Most accounts of moral responsibility claim that it is only appropriately attributed to an agent if she exercises control over the action, omission or consequence for which she is held responsible; it is therefore natural to think that we are responsible for our characters only if we exercise a sufficient degree of the right kind of control over their contents. Accordingly, I devote the first half of the paper to establishing that only the basically virtuous person exercises the requisite control. It is a condition upon responsibility for bringing about a consequence!

e that we are capable of understanding the value of that consequence obtaining or failing to obtain; I show that the vicious are unable to understand this value. I then turn to an examination of various non-control conceptions of responsibility for character; I show that all fail. Responsibility for character requires control, but only those who are already basically virtuous exercise the requisite degree of control.

Are we responsible for our characters, for what we *are*, as well as what we do? On some accounts of responsibility, it is difficult even to make sense of the question. On those theories of moral responsibility descended from Hume, for instance, a person is responsible for her actions just insofar as they *reflect* her character. The character itself is the background against which attributions of responsibility are made, and therefore apparently cannot itself be assessed in this manner. But if we are not responsible for our characters, and our characters are the cause of our actions, then it seems arbitrary to maintain that we are responsible even for our actions. We would like an account of responsibility which is free from such arbitrariness, which goes all the way down, as it were. A growing number of philosophers argue that we can have such an account; that we can legitimately hold people responsible for their characters.<sup>1</sup> In this paper, I will subject their claims to scrutiny. I will argue for a surprising conclusion: the worse our characters, the less we are responsible for them. Good people are responsible for improving their characters; even the mediocre might be properly blamed for their faults. But the truly bad ought to be excused. Or so I shall claim.

In order to fix our ideas, we need some rough definitions of the terms with which we shall be concerned. Unfortunately, both our primary terms, ‘character’, and ‘responsibility’, are used in multiple and sometimes contradictory ways in everyday language. For instance, each word is used both normatively and descriptively; thus we speak of someone’s having no character, when we might equally say that they had a character of which we disapprove, or we might say that someone is not responsible, by which we mean that they are all too likely to be responsible (for mayhem). Somewhat stipulatively, then, I define ‘character’ as:

A person’s characteristic manner of understanding the world, insofar as this sets her apart from other people, and her normal dispositions to actions and emotions, especially as those actions and emotions are morally significant and especially insofar as those dispositions tend to issue in actions.<sup>2</sup>

A satisfactory definition of character will make reference both to characteristic actions, and to dispositions, even when the latter never issue in actions. The ability to overcome desires which are immoral, or simply incompatible with one’s self-image is admirable. But we think worse of a person just for having certain desires, even if they are always kept under control. There is room here for reasonable disagreement about the relative weighting of first- and second-order desires. But we do not need to settle the issue here.

Next, we need a definition of responsibility. Here, we can invoke the well-established usage of the term in moral philosophy. Accordingly, I will say that someone is responsible for something (an action, an emotion, or for their character) if they are the appropriate target for the *reactive attitudes* (indignation, resentment, gratitude, and so forth) with regard to that thing. Here, of course, I follow the well-trodden path opened up by Peter Strawson’s ‘Freedom and Resentment’.<sup>3</sup> The responsibility in question here is, of course, *moral* (and not, say, merely causal) responsibility. Thus, the question with which we shall be concerned might be paraphrased as follows: are we appropriate targets for the reactive attitudes on the basis of our characteristic dispositions to action and emotion?

Put like, this a negative answer might seem intuitively compelling. Following George Sher, we can set out the rationale for this intuitive response as follows:

- (1) no one is responsible for anything beyond their control and
- (2) people (typically) lack control over their characters.<sup>4</sup>

Thus, if we are to conclude that people are indeed responsible for their characters, we shall have to deny either (1) or (2). We shall have to show either that people (typically) possess a

sufficient degree of control over their characters to ground an ascription of responsibility for it, or that they are responsible *despite* a lack of such control. We shall examine both these conditions. Thus, this paper will fall into halves. In the first, we shall examine the extent to which people exercise control over their characters; in the second, we shall turn our attention to arguments which purport to establish that control is not necessary for responsibility.

### ***Responsibility and Control***

Most philosophers believe that control is a necessary condition of moral responsibility. Even in the so-called Frankfurt-type cases, it seems that we are responsible because we exercise control over our actions.<sup>5</sup> Of course, this condition upon the ascription of moral responsibility has been developed and defended primarily with responsibility for actions and omissions in mind; nevertheless, it seems plausible to believe that the same, or a very similar, condition would apply in the case of responsibility for character. In any case, whether or not control is *necessary* for moral responsibility, it is normally sufficient; thus, if we can show that agents typically exercise sufficient control over their characters, we shall have demonstrated their responsibility for them (absent certain excusing conditions). If the contention that agents have such a degree of control is vindicated, then, we shall have no need to investigate condition (1), but will have established that we are responsible for our characters.

However, the contention that we exercise such control over our characters seems implausible. Recall our definition of character:

A person's characteristic manner of understanding the world, insofar as this sets her apart from other people, and her normal dispositions to actions and emotions, especially as those actions and emotions are morally significant and especially insofar as those dispositions tend to issue in actions.

Now, it is not possible *directly* to control (for example) such dispositions. A disposition is not something which we can choose to have or lack; I cannot now choose to be disposed to generous feelings toward someone I have previously resented, or to react calmly to situations which previously caused me to panic. However, lack of direct control does not entail lack of *indirect* control, and it is possible to exercise indirect control over my dispositions. Though I cannot simply choose to have a generous disposition, I can engage in a long-term project aimed at cultivating such a disposition. I can meditate on the benefits of such a disposition for myself and for others, reading inspiring tales of generosity,

associate only with people who themselves display and praise this virtue, and so on. Though there is no guarantee that the sincere and dedicated effort to pursue such a strategy will succeed, many people who engage in this kind of exercise report at least limited success with it.

Thus, though we do not have direct control over our characters, we might be said to have indirect control over them, in the sense that we have the ability to engage in various intentional actions which are designed, and are at least sometimes effective, in altering them. However, it is far from clear that possession of *this* ability is sufficient to demonstrate that we actually exercise the control requisite for moral responsibility. We have this ability – call it ‘raw’ ability – whenever the actions required to bring about a consequence are within our physical power. We possess the raw ability to X when X-ing is something of which we are capable; given sufficient motivation to X, we would do so. Now, possession of the raw ability to X seems insufficient to ground responsibility for failing to X, in a number of clear cases. Consider cases of excusable ignorance, for instance. Many people have died while others stood by helplessly, not knowing that they possessed the raw ability to save the dying person’s life by performing CPR. In some of these cases, this ignorance was excusable, because (for instance) the helpless bystanders lived in cultures in which CPR was not known. No one would accuse these people of being responsible for the deaths of those they watched die. Thus raw ability is not sufficient to ground ascriptions of moral responsibility, and it seems clear that this is the case because raw ability is insufficient for control. I do not have control over the realization of a consequence if I lack the knowledge as to how to bring about (or avoid) that consequence; if my ignorance is not culpable, then I am not responsible for the obtaining of that consequence.

Thus, if we are to establish that I am responsible for my character because it is under my indirect control, we must show that the condition which defeated moral responsibility in the CPR case is not present. The defeating condition is *non-culpable ignorance*: so long as someone non-culpably lacks the knowledge that she possesses the raw ability to X, or (non-culpably) lacks the knowledge that X-ing will or might bring about a certain consequence, Y, she is not responsible for failing to bring about Y. If I am (control) responsible for my character, then, I must know that I am able bring about

changes in it. If I (non-culpably) lack this knowledge, I am not responsible for failing to change it. Thus, raw ability is not sufficient for moral responsibility; in addition, we need to possess the knowledge requisite for exercising that ability (or be culpable for our ignorance).

Moreover, even possession of this ability – call it capability – is insufficient. There are further epistemic conditions upon moral responsibility. In addition to the capability to X, which on our definition includes the knowledge that X is a means of bringing about Y, I must also know that Y is a good which ought to be realized (or, once again, be culpable for my ignorance). Non-culpably lacking this knowledge, I am not responsible for my failure to X, though I possess the capability to do so. Thus, for instance, a person who today suffers from skin cancer due to her over-exposure to the sun as a child might excuse her parents for subjecting her to such exposure on long summer holidays on the grounds that though they possessed the capability to refrain from so exposing her, they were (non-culpably) ignorant as to the effects of the sun, given that the medical opinion of the day emphasized the beneficial effects of sun exposure. Non-culpable ignorance can defeat responsibility when it is ignorance of our ability to X, ignorance of the fact that X-ing brings about Y, or when it is ignorance regarding Y, and its true value.<sup>6</sup>

If I am responsible for my character in virtue of my indirect control over it, then, it must be the case that

- (i) (a) I know that I have indirect control over (significant aspects of) my character, or  
(b) if I lack this knowledge I am culpable for my ignorance, *and*
- (ii) (a) I know what shape my character ought to have or,  
(b) if I lack this knowledge I am culpable for my ignorance.

I must have the capability of shaping my character, and know how my character ought to be shaped.

For present purposes, I shall simply assume that (i) is satisfied, and focus attention on (ii). For it is this condition, I suspect, which defeats responsibility for character (if responsibility requires control), at least so far as people with very bad characters are concerned.

Recall, once again, our definition of character:

A person's characteristic manner of understanding the world, insofar as this sets her apart from other people, and her normal dispositions to actions and emotions, especially as those actions and emotions are morally significant and especially insofar as those dispositions tend to issue in actions.

Now, to a very large extent, character so understood acts as an interpretive framework for us.<sup>7</sup> I see the moral properties of the world *through* my character. Given that this is the case, and given that I have a very bad character, it is very difficult for me to see that my character ought to be different from the way it is, at least in its general outline. Since I can only grasp the morally relevant properties of the world through the interpretive framework of my character, I cannot perceive a discrepancy between the way the world is, morally speaking, and my perception of it.

Contrast, for instance, two people who have an undesirable character trait in common. Let us take the example of a lack of generosity, in the sense that they both tend to suspect people's motives and put the worst possible interpretation upon their behavior. Do they have control over this trait? They both possess the raw ability to alter it (we stipulate). They could read uplifting stories of human sacrifice, review the evidence for genuine altruism, and associate only with people of a generous disposition. They may even have the capability of altering this trait, if they know that they possess this raw ability. But do they have a reason for bringing about this change in themselves? This will depend on how bad their characters are.

Perhaps the biggest obstacle they face to altering this trait is simply coming to be aware of its existence. The trait is, we said, cognitive, in the sense that it functions as an interpretive framework through which the world and other people are perceived. Most of the time, we are not aware of our interpretive frameworks, just as we are not aware of eyeglasses, if we wear them. We look through them, and not at them. Coming to be aware of biases and distortions in our interpretive frameworks, just like our eyeglasses, requires that a discrepancy between what we see and what others see, or what we see at a particular time, and what we cannot help acknowledging at a later time, is forced upon our attention. Given that we cannot take off our interpretive 'spectacles', we might come to see the flaws in them in one of two ways:

- (1) we might be told, by people we trust, that our reactions are distorted, or
- (2) we might (implicitly) make predictions about people's behavior on the basis of our perception of them. When they repeatedly fail to act in the manner predicted, we might come to reconsider our assessment, and, eventually, to question our own ability to make such assessments.

Now, the mediocre ungenerous person is much better placed with regarding to satisfying one or both of these conditions than the truly bad person. Let us take condition (1) first. The mediocre bad character is more likely to associate with – and, just as importantly, trust – people who do not possess this character flaw than the truly bad. The truly bad character is unlikely to have friends much better than she is, since such friends will be repulsed by her badness. Even if she has such friends, she is unlikely to trust their judgment (assuming that she has the ability to trust at all). She is likely to interpret their generous impulses as unrealistic soft-heartedness. She cannot admire their good qualities, since such admiration requires recognition of these qualities *as* good, and therefore requires relatively undistorted moral perception on her part. Given that she has few associates with the ability to point out her flaws, and no reason (as she sees it) to trust any that do, she will not usually be able to satisfy condition (1).<sup>8</sup>

Condition (2) seems, at first sight, less demanding. But we need to be careful in assessing it. The kind of prediction we have in mind here is not like the hypothesis of a scientist, who, under ideal conditions, can specify in advance just what would count as falsifying it. Instead, the prediction will likely be far vaguer, and it will therefore be far less clear whether the behavior in question satisfies it or not. Moreover, the behavior must itself be interpreted, which is to say that it must be seen through the same interpretive framework which generated the prediction in the first place. Once again, the mediocre character will be far better placed to detect such discrepancies than will the truly bad character: the former has a character, and therefore an interpretive framework, which is not uniformly bad, and therefore has a better chance of accurately perceiving her mistake. The truly bad character, on the other hand, has a character which is seamless; accordingly, her moral perception will be all of a piece, and her later interpretations will tend to confirm her predictions.

Thus, it seems that character revision requires character inconsistencies. If there are aspects of our character which are undesirable, we shall be able to detect, and, therefore correct, them only to the extent that our characters are not uniformly bad. Only if our character s are good enough, in parts, to sustain trusting relationships with those who lack our flaws, or good enough to be able (indirectly) to detect flaws in ourselves by becoming aware of discrepancies between what we predict and what takes place, will be able to improve our characters. Without such inconsistencies, each of us necessarily sees ourselves

as the standard of accurate moral perception, against which others are to be measured, and we lack the means to improve our characters.<sup>9</sup>

In other words, though we possess the *capability* of altering our characters, the bad person generally lacks *reasons* for so doing. Unless my character is inconsistent (and I know this), which provides me with a reason to alter aspects of it to integrate them with the whole, or I associate with people who I respect but who regard aspects of my character as undesirable and tell me so, I do not know that I ought to change myself. As we saw, however, meeting condition (ii) requires that I possess this knowledge, or that I am culpable for my ignorance. If we are to rescue the control conception of responsibility for character, we shall therefore need to show that I am culpable for my ignorance of reasons for altering it.

#### *Responsibility and Culpable Ignorance*

According to the standard analyses of the phenomenon, in order for my ignorance to be culpable it must be the result of an act or omission of mine which was itself culpable. In the case of responsibility for character, presumably, this ‘benighting act’ (to adopt Holly Smith’s terminology) would consist in my failure to learn that (aspects of) my character are undesirable or wrong. But for my benighting act to be culpable, it cannot *itself* be performed out of ignorance, or we will be faced with a vicious regress. Culpable ignorance must be traceable back to an initial culpable act that was not, itself, ignorant. The culpably ignorant person is culpable precisely because she knew (at least dispositionally) that her benighting act was wrong, or, as Smith puts it, it must be the case ‘that her unwitting act fell within the risk (known to her) of her benighting act’.<sup>10</sup> Thus, in order for my ignorance of my character’s viciousness to be culpable, this ignorance must be the result of an act or omission of mine that was not ignorant. But there is no reason to believe that there has ever been any such act or omission on the part of the truly bad character. To be sure, she may have been told by various people that her character is bad, but she hears this advice through the interpretive framework *of* her character. The people who told me that I needed to change seemed to me themselves to have faulty characters. ‘They are not realistic’ I might say; ‘they have their heads in the clouds’; ‘they do not understand the ways of the real world’, or another of the phrases which we commonly hear in the mouths of the Thrasymachuses among us. In other words, it will seem to me that those who advise me to

change myself do not have a clear view of moral and practical reality. I might be wrong to so believe, but my ignorance is not the product of an intentional act for which I am culpable.<sup>11</sup>

Some philosophers might object to this account of culpable ignorance on the grounds that it traces responsibility back to a *specific* act or omission. Granted, they might say, there is no such act or omission in the kind of case with which we are concerned; there is no moment at which I chose to develop my character in a specific way, rejecting alternatives. However, this only shows that the account of culpable ignorance needs revision. There are many clear cases of culpable ignorance which do not involve such a specific act, or even an identifiable series of such acts. Consider the case of a doctor who is unable to save the life of a patient because she is unaware of the existence of a new drug. It might or might not be the case that we can trace her ignorance back to some specific act or omission: perhaps she failed to attend an information session organized by the hospital, despite the knowledge that the session was important. But even if there was no such specific omission on her part, she is still culpable for her ignorance. We do not need to trace her ignorance back to some specific act or omission with regard to which she was not ignorant, for doctors have a *generalized duty* to keep up with medical advances. Thus, if she failed to know that there was an opportunity for her to find out about a new technique or drug, that ignorance was itself culpable. When people occupy certain roles, it is not necessary that we can identify a specific benighting act or omission for us to judge that an unwitting act performed as occupant of that role is culpable.<sup>12</sup>

Let us assume that this account of culpable ignorance for the occupants of special roles is correct. Can we not insist that it is precisely this kind of account of culpable ignorance that is appropriate in the case of responsibility for character? Since our moral character is constitutive of our moral identity, and is the locus around which is focused all of our moral acts and omissions, it might be thought that we bear an indefeasible obligation to cultivate it and watch over it. Just as the doctor is obliged to ‘keep up’ with advances in her profession, so each of us is obliged to keep a constant vigil over our characters, engaging in the Socratic enterprise of examining our lives and our souls.

Perhaps this account of our duty with regard to our characters is correct. It will not, however, establish that we are all responsible for our characters. For if we accept the

generalized duty of self-scrutiny, it remains the case that such scrutiny is a coherentist enterprise.<sup>13</sup> If I act conscientiously on this duty, I will examine my character in order to extirpate inconsistencies; to root out those elements which do not cohere with the broad outlines of my character. I will seek out people *I* take to be moral exemplars the better to engage in my Socratic task, read books which will recall me to what *I* take to be the true path. The racist will be a better – more consistent – racist if he accepts this duty (unless he is a racist despite himself); the same will apply to the sexist, the anti-semitic, the religious bigot, and so on. Thus Socratic self-scrutiny will not rescue the vicious person from her viciousness; indeed, it might entrench it further.

Thus, the truly bad person is not responsible for her character, and to the extent to which she seeks, conscientiously, to improve herself, her efforts will only entrench her vices. She may have the capability to alter her character so as to conform more closely to the good, but she lacks reasons for so doing; moreover, she is not culpable for her ignorance. To the extent to which she controls her character, she cannot control it in the light of the good. But the (basically) good person *is* responsible for her character; she really is guided by the good. Any inconsistencies she detects in herself really ought to be corrected, and she is culpable for failing to act on her knowledge. She is responsible for her character, at least on the conception of responsibility upon which we have been working. The basically virtuous person is responsible for making herself more virtuous, while the basically vicious person is not responsible for remaining more vicious; however, neither is responsible for their basic character in its general outlines.<sup>14</sup> Only the good enough are responsible for their character, but they are not responsible for the fact that they are responsible.

### ***Responsibility for Self.***

If we are to show that we are all responsible for our characters, the good, the mediocre and the truly bad alike, we shall have to demonstrate that responsibility does not require control. Given the intuitive plausibility of the control conception of responsibility, expressed in dictum that ought implies can, the burden of proof rests squarely upon those who wish to take this route. I shall examine the arguments of three different philosophers who have attempted to shoulder this burden.

*Trianovksy*

How might we be responsible for our characters, despite our inability to shape their contents to any significant extent? Some philosophers have suggested that while we do not have control over our characters, nevertheless we might have some analogue of control which is sufficient to ground responsibility. Thus, for instance, Gregory Trianovsky agrees that someone is responsible for her characters only if she is responsible for the shape it possesses, but argues that often enough people meet this condition. Despite the fact that it is not up to me to choose the way I am (in the sense that I cannot choose alternatives), the traits I possess might be to my credit or discredit, because *I* am responsible for my having them in the required sense.

How might this be the case? Trianovsky compares two people, ‘Sunny’ and ‘Wimpy’. Both are generous and considerate people, but there is a crucial difference in the manner in which they came to be this way. Wimpy has a very malleable temperament; he would have taken on the values and the dispositions inculcated in him by his parents no matter what they were. Sunny, on the other hand, has a naturally sweet and cheerful temperament, which would have been resistant to being shaped to any form other than that which it actually took. This is a crucial difference between the two: they have much the same character, but Wimpy has it only because he is malleable, whereas Sunny has it because it is in keeping with his natural inclinations. Thus, Wimpy is passive in the formation of his character and Sunny active, in this sense: Sunny’s character is the result of his ‘active discrimination’, rather than (as in Wimpy’s case), the discrimination of others. For this reason, Sunny, and not Wimpy, is responsible for his character.

Trianovsky admits that this account of moral responsibility admits a large element of moral luck into the equation. It will be a matter of luck whether our temperaments are such that we are capable, like Sunny, of responsibility for our characters; indeed, it will be a matter of luck whether the character for which we are responsible will be virtuous or vicious. If I am the unfortunate possessor of a cruel temperament, on Trianovsky’s account I am responsible for my cruel character. This counterintuitive result might already cast some doubt on the account; let me suggest a further reason to reject it.

Trianovsky holds that I am responsible for my character insofar as I possess it as the result of my active discrimination between ends. As he points out, however, these discriminating acts of mine have to be connected in the appropriate manner to the content of the trait which is in the process of being developed. After all, Wimpy can be said to discriminate between traits – he rejects those of which his parents disapprove, and reinforces those of which they approve – but this is insufficient to ground his responsibility for them, since it is not on the basis of the *content* of the traits that he makes these discriminations, but on the basis of something extrinsic to them (his parents' wishes). Thus, as Trianovsky puts it:

“Reference to the value-making features of the trait (or of what it aims at) must figure essentially and fairly directly in the explanation of why the trait is developed and maintained.”<sup>15</sup>

But, when he is acquiring his character, is it really the case that *Sunny* discriminates in the appropriate manner? That is, does Sunny act (say) considerately because he understands the moral reasons why we ought to take other people's interests and welfare into account? Of course not; it is only when Sunny has achieved some degree of moral maturity that he will be able to discriminate *for those reasons*. In the meantime, the reasons why he chooses to act in ways that reinforce these dispositions are extrinsic to the content of the disposition. He does so because it gives him pleasure, not because it is right. Thus Sunny does not meet Trianovsky's conditions on responsibility for character. Nor can anyone, for while we are engaged in the early process of character formation, we are not yet in a position to appreciate the moral reasons that could guide this process. We do not have control over our characters; nor do we possess Trianovskyian quasi-control over their contents.<sup>16</sup>

Thus, if we are to vindicate the suggestion that we are responsible for our characters, it will be in spite of this lack of control. There are, so far as I am aware, two accounts of responsibility for character which hold that control is completely irrelevant; it is to these that I now turn.

#### *Sher*

The first account is due to George Sher. Sher agrees that we are not responsible for our characters, since we lack control over their contents. Nevertheless, he holds that we are *blameworthy* for our undesirable character traits. I propose to treat this as a (non-control) account of responsibility for character, on three grounds:

- 1) Though Sher claims not to be discussing responsibility, his definition of blame seems very close to our definition of moral responsibility. Sher holds that we blame someone when we hold that their faults make them appropriate targets for ‘indignation, disgust, or disappointment’.<sup>17</sup> But these are reactive attitudes. Thus, Sher’s ‘blame’ seems to be equivalent to our ‘moral responsibility’, at least in its negative aspect.
- 2) Sher’s reason for holding that we are not responsible for our characters is that responsibility requires control.<sup>18</sup> But since we are here concerned with the possibility of a non-control conception of responsibility, this contention begs the question against us.
- 3) To the extent that what Sher calls blame *does not* involve the reactive attitudes, ‘blame’ cannot pick out anything interesting at all. If Sher means that though we do not hold people responsible for their characters, nevertheless we have negative attitudes toward them, this is a very uninteresting claim. We have negative attitudes to a great variety of entities we do not hold responsible, including animals and machines. I take it, though, we are interested in attitudes which we appropriately direct at human beings only. But these latter just are the reactive attitudes.

What grounds does Sher offer for holding people responsible for their characters? Essentially, his idea is that we are rightly held responsible for our characters because they reflect well or badly upon us. We are not held responsible for actions over which we exercised no control – perhaps because we were drugged or hypnotized – but this is precisely because these actions do *not* reflect badly or well on us. It is not to my credit if I am compelled to work for Oxfam. It is, however, to my credit if, by chance, I have the disposition to help the needy. A person’s dispositions are bound up with his moral identity; thus:

there is no clear difference between being appalled or disgusted by his dishonesty, manipulateness, and callousness and simply being appalled or disgusted by *him*.<sup>19</sup>

Since I am constituted by my traits, blame for my actions entails blame for the traits they reflect.

However, not all my failings and flaws are appropriate targets of blame, on anyone’s theory. People are not to blame for their stupidity, or for their ugliness, and this despite the fact that stupidity, at least, is constitutive of identity in the same sense as is viciousness. How are we to account for this difference? Sher argues that we are *morally* to blame only for our *moral* faults; since stupidity is not a moral fault, we cannot be morally to blame for it.

How convincing is this account? Remember that we are concerned here with the reactive attitudes. That is, we want to know to what extent it is appropriate to react to people with indignation, resentment, gratitude, and so on, and to what extent we ought

instead to adopt the objective attitude towards them. There are conditions under which the reactive attitudes are not appropriate: we do not, for instance, hold the psychopath or the very young child morally responsible for their actions. Most accounts of these exempting conditions centre around the claim that they prevent the agent from exercising the right kind of control over her actions; thus any account which proposes to dispense with control faces an uphill battle to establish why the reactive attitudes are appropriate. It will not do simply to point to the fact that we react negatively to people on the basis of their character; a negative attitude is not necessarily a reactive attitude. I can be just as disgusted by someone's physical ugliness as their morals; nothing in what Sher says shows that we ought to treat the two cases differently.

There is no doubt, however, that Sher is quite right to point out that it is difficult to draw clear boundaries between someone's actions and her character. To the extent that this is so, and we wish to blame people for their actions, are we not already implicitly committed to blaming them for their character as well? Indeed, on some accounts of holding responsible for action, we do so only to the extent to which these actions reflect the character of the agent.<sup>20</sup> To the extent to which actions are the product of, or reflect, our character, does not responsibility for character stand or fall with that of action?

Though it is true that our actions typically reflect our character, a control account of moral responsibility can be given for our actions (and not for our character). Hence, it is not true that both must be treated in the same manner. It may be that some of our actions really are necessitated by character; this might be the case, for instance, with the psychopath. But it is precisely for this reason that we exempt the psychopath from blame. Sher offers us no reason why we should abandon a control conception of moral responsibility, and therefore no reason why we ought to hold people responsible for their characters.

### *Kupperman*

The final non-control account of responsibility for character we owe to Joel Kupperman. Kupperman's begins by advancing consequentialist considerations in favor of holding people responsible for their characters. We normally praise and blame, Kupperman contends, in order to produce certain effects: principally, to increase the probability that people will conform their actions to correct values. It is because praise and blame are designed to have these effects that we exempt certain people from it in certain situations.

When the praise (blame) cannot be effective in bringing it about that people conform their actions to correct values, because these people are constitutionally unable to appreciate these values, or because their actions do not, for whatever reason, reflect their values, we exempt them; as do we when the blame would be demoralizing, because it would hold someone to values she cannot hope to live by. But neither of these is the case when it comes to praising and blaming people for their characters. People may not be able to change their characters, but they can alter their actions, and praise and blame *for their character* might have this desired effect on their *actions*. Moreover, praise and blame directed at character might have a salutary affect on children, whose characters are not yet fixed; ‘the general pattern of praise and blame of character in a society or in the group within which a child grows up can affect the formation of character’.<sup>21</sup>

Kupperman buttresses these considerations with a burden of proof argument. He holds that, as a matter of fact, we *do* hold people responsible for their characters. The fact that this is so immediately places the burden of proof on those who wish to argue for a revision in our everyday moral discourse. In order to make their case, they will have to demonstrate that our practice of blaming people rests on a mistake. But it doesn’t rest on a mistake, Kupperman argues: the proof is that he has offered a justification of our practice of praising and blaming which shows it to be sensible and effective.

But is the justification he has sketched of (what he take to be) our practice the *right* justification? In order to show that our practice does not rest on a mistake, he needs to do more than show that there is a *possible* rationale for it; he needs to show that the rationale given is in fact the one that underlies the practice. If our practice rests on a different, faulty, rationale, then it *does* rest on a mistake. And this is in fact the case: our practice of holding people responsible for their characters (to the extent to which we have such a practice) does not rest on consequentialist grounds at all, but on the grounds that we are (control) responsible for our characters. Kupperman himself admits as much, by introducing his proposal as a kind of ‘as if’ way of talking. We do not *actually* believe people are responsible for their characters; we simply talk as if we believed it. It is, of course, extremely unlikely that such a fiction underlies our everyday language. If this is the case,

however, Kupperman's suggested justification for this way of talking is beside the point: our practice does rest on a mistake, after all.

And since this is the case, since our practice of holding people responsible for their characters rests on grounds quite different to those Kupperman suggests, then his proposal is (at least) as revisionary as is the alternative I urge. It is at least as revisionary to suggest giving up on holding people responsible for their characters, but continuing to talk as though they were, as it is to suggest the former without the latter. And since this is so, Kupperman's burden of proof argument fails altogether. Since both sides in the debate argue for the introduction of sweeping revisions in everyday discourse, both are revisionary, and neither can rely on the practice as it exists to displace the burden on to its opponent.

Thus, the best arguments of Trianosky, Sher and Kupperman to the contrary, control is necessary for responsibility for character. But, as we have seen, the vicious person does not typically exercise a sufficient degree of control over her character to ground moral responsibility for it. It follows that she is not responsible for her character. Character is malleable, and to the extent to which we really are guided by moral and practical reality, we have a responsibility to alter our characters. But we are not responsible for whether or not we are so guided, whether or not we are responsible. Our character is, as Thomas Nagel rightly argues, a result of our constitutive moral luck;<sup>22</sup> and it is our moral luck which will determine whether we can reshape it in accordance with correct values.

## NOTES

<sup>1</sup> The thesis that we are responsibility for our characters is defended in Robert Merrihew Adams, 'Involuntary Sins' *The Philosophical Review* 94 (January 1985); Michele Moody-Adams, 'On the Old Saw that Character is Destiny', in Owen Flanagan and Amélie Oksenberg Rorty (eds) *Identity, Character and Morality: Essays in Moral Psychology* (Cambridge, Mass.: The MIT Press, 1990); Gregory Trianosky, 'Natural Affection and Responsibility for Character: A Critique of Kantian Views of the Virtues', in Flanagan and Rorty; Joel Kupperman, *Character* (New York: Oxford University Press, 1991); Robert Audi, 'Responsible Action and Virtuous Character' *Ethics* 101 (January 1991); Jonathan Jacobs, *Choosing Character: Responsibility for Virtue and Vice* (Ithaca: Cornell University Press, 2001), and (or so I shall argue), George Sher, 'Blame for Traits' *Noûs* 35 (2001).

<sup>2</sup> This definition draws upon that suggested by Kupperman, p. 17. Like him, I understand 'normal' in a manner relativized to circumstance; thus, it is not a statistical notion, but encompasses the manner in which we typically act in certain (perhaps highly unusual) situations.

<sup>3</sup> Peter Strawson, 'Freedom and Resentment' in John Martin Fischer and Mark Ravizza (eds) *Perspectives on Moral Responsibility* (Ithaca: Cornell University Press, 1993).

<sup>4</sup> Sher, p. 147. I have altered Sher's conditions in several ways; most importantly, I have substituted 'responsibility' for 'blame' and 'characters' for 'traits'. Sher places great weight on the distinction between blame and responsibility; for him, we are to *blame* for our traits, but not *responsible* for them. I shall discuss the manner in which he draws this distinction in the second half of this paper.

<sup>5</sup> Frankfurt-type cases are those in which an agent is responsible for an action which she freely chooses to perform, despite the fact that a counterfactual intervener waits in the wings to ensure that she performs that action. This kind of case was first introduced by Harry Frankfurt, in his 'Alternate Possibilities and Moral Responsibility', reprinted in John Martin Fischer (ed) *Moral Responsibility* (Ithaca: Cornell University Press, 1986). For a defense of the idea that we are responsible for our actions because (and only because) we exercise control over them, even in Frankfurt-type cases, see John Martin Fischer and Mark Ravizza, *Responsibility and Control: A Theory of Moral Responsibility* (Cambridge University Press, 1998).

<sup>6</sup> Thus, *pace* Michele Moody-Adams, the fact that we are *able* (possess the raw ability, or even the capability), to act out of character is insufficient to show that we are responsible *for* our characters. In addition to capability, we need motivating and justifying reasons for so acting.

<sup>7</sup> This view of character is one to which those virtue ethicists who conceive of the virtues (and, presumably, the vices too) as cognitive appear committed. John McDowell's virtue theory is perhaps the best known of the cognitivist variety. See, for instance, his 'Virtue and Reason', in Roger Crisp and Michael Slote (eds) *Virtue Ethics* (Oxford University Press, 1997).

<sup>8</sup> On the way friendship, love and other relationships can enable character change, see Marcia Homiak 'On the Malleability of Character', in Claudia Card (ed) *On Feminist Ethics and Politics* (Lawrence: University Press of Kansas, 1999).

<sup>9</sup> This line of thought is rejected by Robert Audi. Since (he claims) 'one might have a personality antecedently to having a moral character', we have a basis for choosing our character prior to the possession of the interpretive framework it represents (Audi, p. 308). I'm not sure I understand Audi's distinction between personality and character; in any case, the contention will not stand up to scrutiny. We are concerned with *moral* responsibility; if, therefore, I am morally responsible for choosing my character, it will be because I have shaped it in the light of moral reasons which were (or should have been) available to me. Though it is true I might have moral reasons available to me before I have a character, they will not be reasons of *my own*. I will simply have absorbed them, and will not yet be in a position to reflect upon them. After all, we typically exempt very young children from moral responsibility. The ability to reflect upon moral reasons is acquired along with, indeed via, the acquisition of character. But character functions as an interpretive framework; by the time I come to assess my character, the dice are loaded.

<sup>10</sup> Holly Smith, 'Culpable Ignorance', *The Philosophical Review* XCII (October 1983), p. 551. In his 'Moral responsibility and ignorance' (*Ethics*, 107, April 1997), Michael Zimmerman denies that dispositional knowledge of the wrongness of the benighting act is sufficient for culpability. This seems to me to be false, but we need not settle this issue: either account will give the same result in the present case.

<sup>11</sup> Robert Merrihew Adams accepts something like this account of culpable ignorance, yet insists that we are nevertheless responsible for our characters. According to him, we are responsible for our states of mind if these states are intentional and we have them as a result of responding to a set of data which is 'rich enough to permit a fairly adequate ethical appreciation of the state's intentional object' (Adams, p. 26). This, according to Adams, preserves the attractive idea that voluntariness is a criterion of responsibility. However, the question we need to ask ourselves, in ascertaining moral responsibility, is not whether there was rich enough data available – in some sense of that word – to the agent, but whether the agent had the capacity to react to that data. Adams himself excuses 'young children whose experience is not yet rich enough for adequate appreciation of their objects' (26) from moral responsibility; thus he admits that the capacity of the agent to understand the data with which she is presented is crucial to moral responsibility. Excusing children on the grounds they lack this capacity without excusing others who, through no fault of their own, also lack it is entirely arbitrary.

<sup>12</sup> This account of responsibility for the occupants of particular roles is defended by Richard Feldman, 'Voluntary Belief and Epistemic Evaluation', in Matthias Steup (ed), *Knowledge, Truth, and Duty: Essays on Epistemic Justification, Responsibility, and Virtue* (Oxford University Press, 2001). Feldman takes his analysis of role responsibility to show that with regard to some responsibilities, ought does not imply can.

<sup>13</sup> Since reflection on my character is a coherentist enterprise, invoking the image of Neurath's boat (as does Moody-Adams [p. 128]) is perfectly appropriate, but it does not establish that we are responsible for our characters, as she apparently thinks. Why assume that the vicious character is necessarily inconsistent, or that

what inconsistency there is ought rationally – that is, for reasons accessible to the agent, for the ignorance of which he is necessarily culpable – to be settled in favor of the virtues, and not in favor of more consistent viciousness? We need an argument to the effect that (a) the virtues form a unity, (b) vices cannot form a unity, and (c) to the extent that someone has a character that is a mix of vice and virtue, she will always have rationally compelling reasons to resolve the inconsistency in favor of the latter. Provision of such an argument would represent a very large step forward in the rational justification of morality; unfortunately, to my knowledge no one has yet provided anything like it.

<sup>14</sup> Susan Wolf defends a similar asymmetry thesis in *Freedom Within Reason* (New York: Oxford University Press, 1990). Since we are responsible only if we are able to guide our actions in the light of the True and the Good, to the extent to which we are not so able we are not responsible.

<sup>15</sup> Trianovsky, p. 96.

<sup>16</sup> Trianovsky's theory is in some ways similar to that recently developed by Jonathan Jacobs. For Jacobs, we are responsible for our characters because we play an active part in acquiring them. Thus, though the child is closely directed by its parents, and typically absorbs their values, the actions it performs and fails to perform as its character is developed are performed voluntarily. Thus, Jacobs holds, though the agent's character is the result of her circumstances and her temperament, for neither of which she is responsible, nevertheless the fact that she acts voluntarily is sufficient to ground responsibility:

Granted, with a child or even an adolescent this voluntary behavior often involves action directed by another without the young person having initiated the act in the sense of having originated the thought on one's own to do it. Juvenile actions take place in a context of norms, reprimands, expectations, and inducements almost wholly shaped by others. In addition, the young person may not yet be able or inclined to imagine different ways of acting. Yet, even these constraints do not close off all space for voluntariness (15).

But surely it is counterintuitive to hold the adult responsible for the character the child has made, a character which, as Jacobs himself admits, may well be 'ethically disabled' in the sense that it is constitutively unable to appreciate right ethical considerations (34), when we do not hold the *child* responsible for her acts and omissions, precisely because *she* is unable to appreciate the relevant ethical considerations? Voluntariness, in Jacob's sense, is insufficient to ground moral responsibility, or we would hold the brainwashed, the psychopathic and the very young responsible.

<sup>17</sup> Sher, p. 155.

<sup>18</sup> Ibid. pp. 146-7.

<sup>19</sup> Ibid, p. 155.

<sup>20</sup> Edward Sankowski defends such a (roughly, Humean) account in 'Freedom, Determinism and Character' *Mind* 89 (Jan 1980), pp. 106-113.

<sup>21</sup> Kupperman, p. 62

<sup>22</sup> In 'Moral Luck' in *Mortal Question* (Cambridge University Press, 1979), p. 28.