

The Impossibility of Moral Responsibility

Galen Strawson

I

There is an argument, which I will call the Basic Argument, which appears to prove that we cannot be truly or ultimately morally responsible for our actions. According to the Basic Argument, it makes no difference whether determinism is true or false. We cannot be truly or ultimately morally responsible for our actions in either case.

The Basic Argument has various expressions in the literature of free will, and its central idea can be quickly conveyed. (1) Nothing can be *causa sui* – nothing can be the cause of itself. (2) In order to be truly morally responsible for one's actions one would have to be *causa sui*, at least in certain crucial mental respects. (3) Therefore nothing can be truly morally responsible.

In this paper I want to reconsider the Basic Argument, in the hope that anyone who thinks that we can be truly or ultimately morally responsible for our actions will be prepared to say exactly what is wrong with it. I think that the point that it has to make is obvious, and that it has been underrated in recent discussion of free will – perhaps because it admits of no answer. I suspect that it is obvious in such a way that insisting on it too much is likely to make

it seem less obvious than it is, given the innate contrasuggestibility of human beings in general and philosophers in particular. But I am not worried about making it seem less obvious than it is so long as it gets adequate attention. As far as its validity is concerned, it can look after itself.

[...]

[E]ssentially the same argument can be given in a more natural form. (1) It is undeniable that one is the way one is, initially, as a result of heredity and early experience, and it is undeniable that these are things for which one cannot be held to be in any way responsible (morally or otherwise). (2) One cannot at any later stage of life hope to accede to true moral responsibility for the way one is by trying to change the way one already is as a result of heredity and previous experience. For (3) both the particular way in which one is moved to try to change oneself, and the degree of one's success in one's attempt at change, will be determined by how one already is as a result of heredity and previous experience. And (4) any further changes that one can bring about only after one has brought about certain initial changes will in turn be determined, via the initial changes, by heredity and previous experience. (5) This may not be the whole story, for it may be that some changes in the way one is are traceable not to heredity and experience but to the influence of indeterministic or random factors. But it is absurd to suppose that indeterministic or random factors, for which one is *ex hypothesi* in no way responsible, can in

Galen Strawson, "The Impossibility of Moral Responsibility," originally from *Philosophical Studies*, 75 (1994), 5, 7–10, 13–15, 25. Updated and reprinted in *Real Materialism and Other Essays* (Oxford University Press, 2008). Reprinted with kind permission of the author.

themselves contribute in any way to one's being truly morally responsible for how one is.

The claim, then, is not that people cannot change the way they are. They can, in certain respects (which tend to be exaggerated by North Americans and underestimated, perhaps, by Europeans). The claim is only that people cannot be supposed to change themselves in such a way as to be or become truly or ultimately morally responsible for the way they are, and hence for their actions.

II

I have encountered two main reactions to the Basic Argument. On the one hand it convinces almost all the students with whom I have discussed the topic of free will and moral responsibility. On the other hand it often tends to be dismissed, in contemporary discussion of free will and moral responsibility, as wrong, or irrelevant, or fatuous, or too rapid, or an expression of metaphysical megalomania.

I think that the Basic Argument is certainly valid in showing that we cannot be morally responsible in the way that many suppose. And I think that it is the natural light, not fear, that has convinced the students I have taught that this is so. That is why it seems worthwhile to restate the argument in a slightly different – simpler and looser – version, and to ask again what is wrong with it.

Some may say that there is nothing wrong with it, but that it is not very interesting, and not very central to the free will debate. I doubt whether any non-philosopher or beginner in philosophy would agree with this view. If one wants to think about free will and moral responsibility, consideration of some version of the Basic Argument is an over-whelmingly natural place to start. It certainly has to be considered at some point in a full discussion of free will and moral responsibility, even if the point it has to make is obvious. Belief in the kind of absolute moral responsibility that it shows to be impossible has for a long time been central to the Western religious, moral, and cultural tradition, even if it is now slightly on the wane (a disputable view). It is a matter of historical fact that concern about moral responsibility has been the main motor... of discussion of the issue of free will. The only way in which one might hope to show (1) that the Basic Argument was not central to the free will debate

would be to show (2) that the issue of moral responsibility was not central to the free will debate. There are, obviously, ways of taking the word 'free' in which (2) can be maintained. But (2) is clearly false none the less.

In saying that the notion of moral responsibility criticized by the Basic Argument is central to the Western tradition, I am not suggesting that it is some artificial and local Judaeo-Christian-Kantian construct that is found nowhere else in the history of the peoples of the world, although even if it were that would hardly diminish its interest and importance for us. It is natural to suppose that Aristotle also subscribed to it, and it is significant that anthropologists have suggested that most human societies can be classified either as 'guilt cultures' or as 'shame cultures'. It is true that neither of these two fundamental moral emotions necessarily presupposes a conception of oneself as truly morally responsible for what one has done. But the fact that both are widespread does at least suggest that a conception of moral responsibility similar to our own is a natural part of the human moral-conceptual repertoire.

In fact the notion of moral responsibility connects more tightly with the notion of guilt than with the notion of shame. In many cultures shame can attach to one because of what some member of one's family – or government – has done, and not because of anything one has done oneself; and in such cases the feeling of shame need not (although it may) involve some obscure, irrational feeling that one is somehow responsible for the behaviour of one's family or government. The case of guilt is less clear. There is no doubt that people can feel guilty (or can believe that they feel guilty) about things for which they are not responsible, let alone morally responsible. But it is much less obvious that they can do this without any sense or belief that they are in fact responsible.

III

Such complications are typical of moral psychology, and they show that it is important to try to be precise about what sort of responsibility is under discussion. What sort of 'true' moral responsibility is being said to be both impossible and widely believed in?

An old story is very helpful in clarifying this question. This is the story of heaven and hell. As I understand it, true moral responsibility is responsibility

of such a kind that, if we have it, then it *makes sense*, at least, to suppose that it could be just to punish some of us with (eternal) torment in hell and reward others with (eternal) bliss in heaven. The stress on the words 'makes sense' is important, for one certainly does not have to believe in any version of the story of heaven and hell in order to understand the notion of true moral responsibility that it is being used to illustrate. Nor does one have to believe in any version of the story of heaven and hell in order to believe in the existence of true moral responsibility. On the contrary: many atheists have believed in the existence of true moral responsibility. The story of heaven and hell is useful simply because it illustrates, in a peculiarly vivid way, the *kind* of absolute or ultimate accountability or responsibility that many have supposed themselves to have, and that many do still suppose themselves to have. It very clearly expresses its scope and force.

But one does not have to refer to religious faith in order to describe the sorts of everyday situation that are perhaps primarily influential in giving rise to our belief in true responsibility. Suppose you set off for a shop on the evening of a national holiday, intending to buy a cake with your last ten pound note. On the steps of the shop someone is shaking an Oxfam tin. You stop, and it seems completely clear to you that it is entirely up to you what you do next. That is, it seems to you that you are truly, radically free to choose, in such a way that you will be ultimately morally responsible for whatever you do choose. Even if you believe that determinism is true, and that you will in five minutes time be able to look back and say that what you did was determined, this does not seem to undermine your sense of the absoluteness and inescapability of your freedom, and of your moral responsibility for your choice. The same seems to be true even if you accept the validity of the Basic Argument stated in section I, which concludes that one cannot be in any way ultimately responsible for the way one is and decides. In both cases, it remains true that as one stands there, one's freedom and true moral responsibility seem obvious and absolute to one.

Large and small, morally significant or morally neutral, such situations of choice occur regularly in human life. I think they lie at the heart of the experience of freedom and moral responsibility. They are the fundamental source of our inability to give up belief in true or ultimate moral responsibility. There are further questions to be asked about why human beings experience these situations of choice as they do.

It is an interesting question whether any cognitively sophisticated, rational, self-conscious agent must experience situations of choice in this way. But they are the experiential rock on which the belief in true moral responsibility is founded.

[...]

IV

Let me now restate the Basic Argument in very loose – as it were conversational – terms. New forms of words allow for new forms of objection, but they may be helpful none the less.

- (1) You do what you do, in any situation in which you find yourself, because of the way you are.

So

- (2) To be truly morally responsible for what you do you must be truly responsible for the way you are – at least in certain crucial mental respects.

Or:

- (1) What you intentionally do, given the circumstances in which you (believe you) find yourself, flows necessarily from how you are.

Hence

- (2) you have to get to have some responsibility for how you are in order to get to have some responsibility for what you intentionally do, given the circumstances in which you (believe you) find yourself.

Comment: Once again the qualification about 'certain mental respects' is one I will take for granted. Obviously one is not responsible for one's sex, one's basic body pattern, one's height, and so on. But if one were not responsible for anything about oneself, how one could be responsible for what one did, given the truth of (1)? This is the fundamental question, and it seems clear that if one is going to be responsible for any aspect of oneself, it had better be some aspect of one's mental nature.

I take it that (1) is incontrovertible, and that it is (2) that must be resisted. For if (1) and (2) are conceded

the case seems lost, because the full argument runs as follows.

- (1) You do what you do because of the way you are.

So

- (2) To be truly morally responsible for what you do you must be truly responsible for the way you are – at least in certain crucial mental respects.

But

- (3) You cannot be truly responsible for the way you are, so you cannot be truly responsible for what you do.

Why can't you be truly responsible for the way you are? Because

- (4) To be truly responsible for the way you are, you must have intentionally brought it about that you are the way you are, and this is impossible.

Why is it impossible? Well, suppose it is not. Suppose that

- (5) You have somehow intentionally brought it about that you are the way you now are, and that you have brought this about in such a way that you can now be said to be truly responsible for being the way you are now.

For this to be true

- (6) You must already have had a certain nature N in the light of which you intentionally brought it about that you are as you now are.

But then

- (7) For it to be true that you and you alone are truly responsible for how you now are, you must be truly responsible for having had the nature N in the light of which you intentionally brought it about that you are the way you now are.

So

- (8) You must have intentionally brought it about that you had that nature N, in which case you must have existed already with a prior nature in the light of which you intentionally brought it about that you had the nature N in the light of which you intentionally brought it about that you are the way you now are ...

Here one is setting off on the regress. Nothing can be *causa sui* in the required way. Even if such causal 'aseity' is allowed to belong unintelligibly to God, it cannot be plausibly be supposed to be possessed by ordinary finite human beings. 'The *causa sui* is the best self-contradiction that has been conceived so far', as Nietzsche remarked in 1886:

it is a sort of rape and perversion of logic. But the extravagant pride of man has managed to entangle itself profoundly and frightfully with just this nonsense. The desire for 'freedom of the will' in the superlative metaphysical sense, which still holds sway, unfortunately, in the minds of the half-educated; the desire to bear the entire and ultimate responsibility for one's actions oneself, and to absolve God, the world, ancestors, chance, and society involves nothing less than to be precisely this *causa sui* and, with more than Baron Münchhausen's audacity, to pull oneself up into existence by the hair, out of the swamps of nothingness ... (*Beyond Good and Evil*, § 21).

The rephrased argument is essentially exactly the same as before, although the first two steps are now more simply stated. It may seem pointless to repeat it, but the questions remain. Can the Basic Argument simply be dismissed? It is really of no importance in the discussion of free will and moral responsibility? (No and No) Shouldn't any serious defense of free will and moral responsibility thoroughly acknowledge the respect in which the Basic Argument is valid before going on to try to give its own positive account of the nature of free will and moral responsibility? Doesn't the argument go to the heart of things if the heart of the free will debate is a concern about whether we can be truly morally responsible in the absolute way that we ordinarily suppose? (Yes and Yes)

We are what we are, and we cannot be thought to have made ourselves *in such a way* that we can be held to be free in our actions *in such a way* that we can be held to be morally responsible for our actions *in such a way* that any punishment or reward for our actions is

ultimately just or fair. Punishments and rewards may seem deeply appropriate or intrinsically 'fitting' to us in spite of this argument, and many of the various institutions of punishment and reward in human society appear to be practically indispensable in both their legal and non-legal forms. But if one takes the notion of justice that is central to our intellectual and cultural tradition seriously, then the evident consequence of the Basic Argument is that there is a fundamental sense in which no punishment or reward is ever ultimately just. It is exactly as just to punish or reward people for their actions as it is to punish or reward them for the (natural) colour of their hair or the (natural) shape of their faces.
[...]

V

There is nothing new in the somewhat incantatory argument of this paper. It restates certain points that

may be in need of restatement. 'Everything has been said before', said André Gide, echoing La Bruyère, 'but since nobody listens we have to keep going back and beginning all over again.' This is an exaggeration, but it may not be a gross exaggeration, so far as general observations about the human condition are concerned.

The present claim, in any case, is simply this: time would be saved, and a great deal of readily available clarity would be introduced into the discussion of the nature of moral responsibility, if the simple point that is established by the Basic Argument were more generally acknowledged and clearly stated. Nietzsche thought that thorough-going acknowledgement of the point was long overdue, and his belief that there might be moral advantages in such an acknowledgement may deserve further consideration.