

---

# ETHICS AND FREE WILL

## VARIETIES OF RESPONSIBILITY

T.M. SCANLON\*

At least three kinds of “responsibility” may seem to be threatened by the possibility that all of our actions and attitudes are caused by factors outside us, over which we have no control: *personal* responsibility, *moral* responsibility, and *substantive* responsibility.<sup>1</sup>

*Personal* responsibility is threatened by the possibility that all our actions and attitudes are caused by factors outside us, if this would mean that our actions would not be “ours” in the sense required for them to have significance for us. We might seem to be mere puppets, or prisoners, of our circumstances.<sup>2</sup>

*Moral* responsibility is threatened by the possibility that our actions and attitudes are caused by factors outside of us, if this would mean that we could not properly be praised or blamed for these actions or attitudes, because they would not be “up to us” in the required sense.

*Substantive* responsibility is threatened by the possibility that all our actions are caused by factors outside of us over which we have no control, if this would mean that we would lack the kind of control over these actions that is required if they are to affect our moral obligations to others, and others’ obligations to us. Our promises might not be binding because they were not entered into freely. Our decisions to forego certain benefits, or to take certain

---

\* Alford Professor of Natural Religion, Moral Philosophy, and Civil Polity, Harvard University, Department of Philosophy.

<sup>1</sup> I have explored these distinctions in T.M. SCANLON, *MORAL DIMENSIONS: PERMISSIBILITY, MEANING, BLAME* 198-204 (2008) (drawing a distinction between two types of moral responsibility: the type to which moral praise and blame attach, and substantive obligations to others); T.M. Scanlon, *The Significance of Choice*, in 8 *THE TANNER LECTURES ON HUMAN VALUES* 149, 152-58 (Sterling M. McMurrin ed., 1988) (describing the importance of “choice” to personal responsibility, moral blameworthiness, and substantive responsibility).

<sup>2</sup> This threat is discussed in ROBERT KANE, *THE SIGNIFICANCE OF FREE WILL* 79-101 (1996) (discussing the “significance” of an incompatibilist conception of free will to our creativity, autonomy, and understanding of desert for one’s achievements).

risks, may not reduce the obligations of others to help us because these decisions were not made voluntarily.<sup>3</sup>

Each of the last two threats might rightly be called a threat to *moral* responsibility, but they seem to me to be distinct. To separate them, I reserve the term *moral* responsibility for the former – the one concerned with blame. I refer to the latter as a threat to *substantive* responsibility because it has to do with the content of our obligations.

A number of different responses to these threats are mentioned in *Justice for Hedgehogs*. One response is that we cannot help seeing ourselves as actually making choices, and cannot help assigning significance to them, even moral significance.<sup>4</sup> The claim is that it is not possible that theoretical beliefs about the causes of our actions could undermine this significance.<sup>5</sup> A second response is that the best explanation of cases in which it is generally recognized that agents are not responsible for their behavior – such as actions done in ignorance or under hypnosis, and actions by young children – is that this is simply not the case.<sup>6</sup> These agents are causally responsible for their actions, but their actions might be morally excusable or justifiable nonetheless because of external factors.<sup>7</sup>

Neither of these responses is by itself compelling. As to the first, even if it is true that when we are engaging in the process of deliberating and deciding we cannot help seeing this process as something that makes a difference and has moral and ethical significance, it is not clear what follows from this. It might simply be the case that this is an unshakable illusion. As for the second response, it seems to me quite clear that outside causes are not the best explanation for most excusing or justifying conditions. But the idea that such causes undermine moral responsibility may have other sources.

Whatever these sources may be, however, if blame requires a kind of freedom distinct from the kind of freedom required by other forms of evaluation – such as an assessment of artistic talent – there must be something about the nature of blame itself that explains this fact. Thus, a satisfactory resolution of this issue must begin with an adequate characterization of blame, from which we can see why blame requires a certain form of freedom, if indeed it does. More generally, a satisfactory resolution of any of the three

---

<sup>3</sup> This threat is particularly important for Dworkin because his view of equality places great weight on the distinction between inequalities resulting from the choices of the disadvantaged person and inequalities unavoidable due to innate differences in talent, genetic luck, etc. See, e.g., RONALD DWORKIN, *JUSTICE FOR HEDGEHOGS* (forthcoming 2010) (Apr. 17, 2009 manuscript at 221-25, on file with the Boston University Law Review); Ronald Dworkin, *What Is Equality? Part 2: Equality of Resources*, 10 PHIL. & PUB. AFF. 283, 312-13 (1981).

<sup>4</sup> DWORKIN, *supra* note 3 (manuscript at 140).

<sup>5</sup> See *id.*

<sup>6</sup> *Id.* (manuscript at 150-51).

<sup>7</sup> *Id.*

---

---

problems I have listed must start from an account of the values involved – an account of the kind of significance that our own actions and attitudes have for us, an account of the nature of moral blame, or an account of how substantive obligations are created and transferred by voluntary actions – on the basis of which it can be seen whether or not these forms of significance require some form of freedom. What incompatibilists seek to show is that it follows from the best understanding of these values that the forms of significance in question require a kind of freedom that we do not have if our actions are caused by factors outside us, over which we have no control. What compatibilists seek to show is that according to the best understanding of these values, this is not the case.<sup>8</sup> This latter interpretive task is undertaken in *Justice for Hedgehogs*.<sup>9</sup>

Of the three kinds of responsibility I have mentioned – *personal* responsibility, *moral* responsibility, and *substantive* responsibility – the first appears to be the most fundamental. Whatever the exact nature of blame, it would seem to depend in large part on the attitudes expressed in an action. Further, there would seem to be a close relation between the conditions that make an attitude “belong to an agent” in the way required to make it significant for him or her and the conditions required to make an attitude “belong to an agent” in the way required to make it significant for others. Similarly, it would seem that the reasons why an agent’s obligations are sensitive to his or her choices would depend on the reasons for those choices being significant for the agent. So, I will focus initially on the first of my three cases, the kind of significance I call *personal*. This is in line with the strategy of *Justice for Hedgehogs*, which is to start with ethics – what it is for an individual to live a good life – rather than with morality – what we owe to others.<sup>10</sup>

According to *Justice for Hedgehogs*, there are two “ethical principles about how to live not, just in themselves, moral principles about how to treat other people”<sup>11</sup>:

I shall . . . call the first the principle of self-respect. Each person must take his own life seriously: he must accept that it is a matter of importance that his life be a successful performance rather than a wasted opportunity. I shall call the second the principle of authenticity. Each

---

<sup>8</sup> See, e.g., *id.* (manuscript at 139) (describing both compatibilists and incompatibilists). As I am using the term, an incompatibilist is someone who holds that one or more of the forms of responsibility I have mentioned (typically, it is moral responsibility that is emphasized) is ruled out if our actions are caused by factors outside us, over which we have no control. A compatibilist is someone who holds that this is not the case: that responsibility of the relevant kind is compatible with causation by outside factors.

<sup>9</sup> See *id.* (manuscript at 169) (arguing that the concept of responsibility is entirely consistent with a causally determined world).

<sup>10</sup> *Id.* (manuscript at 121).

<sup>11</sup> *Id.* (manuscript at 128).

---

---

person has a special, personal responsibility for identifying what counts as success in his own life; he has a personal responsibility to create that life through a coherent narrative that he himself has chosen and endorses. Together the two principles offer a conception of human dignity: dignity requires self-respect and authenticity.<sup>12</sup>

Authenticity, Dworkin continues, “is an ethical concept: it makes two demands. First, it assigns each person a non-delegable responsibility for identifying and then pursuing his own conception of what it is to live well, to give structure to his life through values he endorses that sustain a coherent narrative of living.”<sup>13</sup> However, it also demands another kind of responsibility:

The second principle of dignity demands . . . that I accept judgmental and liability responsibility in appropriate circumstances. People who blame their parents or other people or society at large for their own mistakes, or who cite some form of genetic determinism to absolve themselves of any responsibility for how they have acted, lack dignity because dignity requires owing up to what one has done. “The buck stops here” is an important piece of ethical wisdom. The second principle also requires taking responsibility in a different, more material, way: dignity requires that I not expect others to subsidize my decisions by bearing their financial or other costs. I do not take responsibility for my own life if I demand that others absorb the cost of my choices: living well means making choices, and that means choosing a life with an eye to the consequences of that life that I should bear myself.<sup>14</sup>

Responsibility, in this latter sense requires two capacities:

First, to be responsible people must have some minimal ability to form true beliefs about the world, about the mental states of other people, and about the likely consequences of what they do. Someone who is unable to grasp the fact that guns can harm people is not responsible when he shoots someone. Second, people must have, to a normal level, the ability to make decisions that fit what we might call the agent’s normative personality: his desires, preferences, convictions, attachments, loyalties and self-image. Genuine decisions, we think, are purposive, and someone who cannot match his final decisions to any of his desires, plans, convictions or attachments is incapable of responsible action.<sup>15</sup>

In *Justice for Hedgehogs*, these theses are put forward as ethical *duties* – principles we *must* follow – and as claims about what is required in order to be an agent who is responsible for his or her attitudes and actions. For present purposes, however, I want to consider them in a slightly different light, as an account of the kind of life that a person has reason to want to live, in which his

---

<sup>12</sup> *Id.*

<sup>13</sup> *Id.* (manuscript at 132).

<sup>14</sup> *Id.* (manuscript at 132-33).

<sup>15</sup> *Id.* (manuscript at 143).

---

---

or her actions and attitudes will have the distinctive personal significance of being his or hers. We can then take this account as a basis for considering what kind of freedom is required in order for an agent to have such a life.

Put in this way, the thesis of *Justice for Hedgehogs* is this: what a person has reason to want is an authentic life – one that accords with and reflects the decisions he or she makes about how to live, and in which these decisions themselves reflect his or her desires, plans, and convictions and are based on true beliefs about the world. The crucial thing about this conception is that – apart from the part about true belief – requirements of an authentic life have to do only with the *relations among* an agent’s mental states – his or her plans, desires, convictions, decisions, beliefs, intentions, and so on – and the actions that flow from them. Dworkin is not concerned about the causal antecedents to one’s attitudes, beliefs, etc.<sup>16</sup> His thesis is that as long as an agent’s life has the required shape, the mental events that are parts of it belong to him or her in the full sense required for them to be meaningful and significant.

In particular, the significance of our decisions to us does not require that they not be caused by factors outside us. But why not? Two reasons seem to be offered, so closely related that they are really just positive and negative sides of the same point. Put positively, the point is that what gives decisions their significance for us – what makes them ours – is that they are reflective of our felt desires, plans, convictions, and so on – that is, of our actual psychological states.<sup>17</sup> Put negatively, it is that decisions that were independent of our actual, present psychology would be based on nothing that would give them meaning for us, or make them ours in any meaningful sense.<sup>18</sup> A “self” detached from our actual, empirical psychology, would be empty.

I find these points entirely convincing.<sup>19</sup> However, we should consider why they might be resisted. It can be agreed, I think, that whether our actions are caused by outside factors or not, any decision that is meaningful for an agent will have to be based on current values, plans, and desires of the agent that, at the moment of this decision, are not being questioned but are held fixed (although they might be called into question and reevaluated at another time). Decision and reevaluation is always piecemeal.

What an incompatibilist would see as a threat is the possibility that these processes of deciding and reevaluating are themselves determined by factors outside of us, over which we have no control. Why should this be a threat to

---

<sup>16</sup> *See id.*

<sup>17</sup> *Id.* (manuscript at 137-61).

<sup>18</sup> *Id.*

<sup>19</sup> Also convincing, but more controversial, is the view, which is part of Dworkin’s thesis, that the crucial relations between an agent’s decisions need not be *causes* of his or her subsequent actions and other mental states. *Id.* (manuscript at 157). They could as well be epiphenomenal. *Id.* (manuscript at 146, 157). What is crucial is that the underlying neural causal mechanisms, whatever they may be, are such as to produce the relevant conformity between desires, beliefs, decisions, plans, and actions.

---

---

the meaningfulness, for us, of our attitudes? It might be because these outside factors could force us – so to speak – to reach certain conclusions, overriding our own judgment. But, as Hume pointed out long ago, this is a mistake. The factors in question determine what our judgment is; they do not override it.<sup>20</sup> A second way in which outside causes might threaten the meaningfulness of our decision-making process would be if the effectiveness of these causes indicates that our normative judgments are influenced or even controlled by factors that are normatively irrelevant. Yet, this does not follow either. The fact – if it is a fact – that our mathematical judgments are determined by factors outside of us does not mean that these judgments are affected by factors that are irrelevant to the mathematical correctness of the conclusions we reach, and there is no reason to think that this is necessarily so in other cases.

Even if the account of ethical responsibility I have sketched is adequate, however, the question remains whether the conditions that *Justice for Hedgehogs* holds to be sufficient for ethical responsibility (what I call *personal responsibility*) are also all that is required in order for it to be appropriate to blame agents for their actions and choices, and all that is required to make those choices significant in altering an agent's obligations to others. In order to answer this question, as I said, we need to have some account of what blame involves, and how obligations are determined. *Justice for Hedgehogs* does not go into much detail about these issues. I want to elaborate, however, on one claim that it does make and to consider some implications of this elaboration that may depart somewhat from the book. The claim I have in mind is in a passage I quoted earlier:

The second principle of dignity demands . . . that I accept judgmental and liability responsibility in appropriate circumstances. People who blame their parents or other people or society at large for their own mistakes, or who cite some form of genetic determinism to absolve themselves of any responsibility for how they have acted, lack dignity because dignity requires owing up to what one has done. “The buck stops here” is an important piece of ethical wisdom. The second principle also requires taking responsibility in a different, more material, way: dignity requires that I not expect others to subsidize my decisions by bearing their financial or other costs. I do not take responsibility for my own life if I demand that others absorb the cost of my choices: living well means making choices, and that means choosing a life with an eye to the consequences of that life that I should bear myself.<sup>21</sup>

The underlying point here seems to me correct and important. Insofar as an individual regards her actions under certain circumstances as belonging to her in the sense required to be meaningful for her life, she cannot deny that these actions are significant for other people's reactions to her, and for their relations

---

<sup>20</sup> DAVID HUME, A TREATISE OF HUMAN NATURE, bk. 2, pt. 3, § 1, at 407 (L.A. Selby-Bigge ed., Oxford Univ. Press 1960) (1739).

<sup>21</sup> DWORKIN, *supra* note 3 (manuscript at 132-33).

with her; nor can she deny that they have implications for her obligations to them – at least not on the ground that they are not hers in the required sense. But from the fact that she cannot object to blame, or to obligations that are held to follow from her actions, on this particular ground, nothing follows about whether she has to accept responsibility for any particular action in either the moral or substantive sense. The fact that any attitude that occurs in the mental life of a person who has the two capacities – the “ability to form true beliefs about the world” and the “ability to make decisions that fit what we might call the agent’s normative personality”<sup>22</sup> – belongs to her still leaves open questions of blame and the shifting of responsibilities.

In one sense, this is obvious. Good actions as well as bad actions are attributable to an agent; but the fact that an action belongs to an agent leaves open the question whether blame or some other attitude is appropriate.<sup>23</sup> The point I have in mind goes beyond this.

According to *Justice for Hedgehogs*, a person is in control of his or her actions in the way required for moral responsibility if he or she possesses the two capacities I have described.<sup>24</sup> Dworkin remarks that this idea of control is very different from the one underlying the incompatibilist’s understanding, where the question is whether an external force was strong enough to overcome the will’s “normal causal role.”<sup>25</sup> But this may not make clear how different the two conceptions are, insofar as it leaves the impression that they are simply two different conceptions of the kind of control an agent must have over his or her attitudes in order for them to belong to him or her in the relevant sense. The more radical position I have in mind (which may or may not be Dworkin’s) is that the two capacities – although they are conditions required in order for an agent to be the kind of being toward whom we can stand in the relations that make blame and obligation make sense – are not conditions whose fulfillment in every given case makes an attitude or action attributable to an agent.

Under this position, having the capacities that Dworkin mentions *to a normal level*<sup>26</sup> does not mean that they always work. Further, in order for attitudes to belong to an agent in the sense required for blame, they need not be ones that the agent could control through the use of these general capacities. A

---

<sup>22</sup> *Id.* (manuscript at 143).

<sup>23</sup> *See id.* (manuscript at 158-60) (suggesting, in the context of discussing the insanity and duress defenses, that reasonable people can disagree about whether someone is truly responsible, and thus deserves punishment, for his or her actions).

<sup>24</sup> *Id.* (manuscript at 158).

<sup>25</sup> *Id.* (“When someone claims that he committed his criminal act in a blind rage or when overcome by an irresistible impulse or under extreme duress of some kind . . . we would ask: were these forces or influences strong enough, in the circumstances, so that they displaced his will’s normal causal role like a drunken sailor pushing the helmsman aside and taking the wheel?”).

<sup>26</sup> *Id.* (manuscript at 143).

---

---

person can be blamed for attitudes and actions that were, or would have been, quite resistant to her better judgment. The fact that she had these contrary judgments would complicate the picture of what she is like, and could therefore modify the kind of blame that is appropriate, but it does not change the fact that the recalcitrant attitudes belong to the agent – they are part of what he or she is like.

If an agent has the two capacities that Dworkin describes, then every attitude that occurs in his or her mental life belongs to him or her in the relevant sense, whether or not the attitude could be controlled by the exercise of those capacities.<sup>27</sup> Conditions that are often said to modify responsibility, such as ignorance of fact, duress, and even, I would say, many forms of mental illness, do not have this effect because they make the action one that does not belong to the person. Instead, what they do is to affect what an action or attitude that belongs to the person indicates about him or her. By this, I mean not what it indicates about *him or her* as a disembodied self, but what it indicates about that person's actual overall empirical psychology – about what he or she will feel and do on other occasions. An action performed in ignorance, or under duress belongs to the agent in the full sense; under some description, it is his or her intentional action, performed for some reason. The question is, what does the action tell us about the agent more generally? In some cases, such as hypnosis or brain stimulation, perhaps, the answer is that the action tells us nothing: any agent would have done the same in response to this treatment. But in most cases, the answer will be a matter of degree, or rather of difference. Willingness to perform an action that causes injury when one is under the influence of a false belief about its consequences is something different from willingness to do so with full knowledge of its effects. Willingness to cause harm under duress indicates something different from willingness to do so for fun or for profit.

I am not certain that this is the position that Dworkin means to hold. But if it is, this makes clear why in his view it is the relations *among* an agent's attitudes, rather than their causal antecedents more generally, that is crucial for responsibility.

---

<sup>27</sup> I defended this view in T.M. Scanlon, *Reasons and Passions*, in *CONTOURS OF AGENCY: ESSAYS ON THEMES FROM HARRY FRANKFURT* 165 (Sarah Buss & Lee Overton eds., 2002).