



# On the Parfitian Thesis of Moral Responsibility

**DOI:**

[10.1007/s11406-018-9999-0](https://doi.org/10.1007/s11406-018-9999-0)

**Document Version**

Accepted author manuscript

[Link to publication record in Manchester Research Explorer](#)

**Citation for published version (APA):**

MirzaeiGhazi, S. (2019). On the Parfitian Thesis of Moral Responsibility. *Philosophia*, 47(3), 587-600. Article 2. <https://doi.org/10.1007/s11406-018-9999-0>

**Published in:**

Philosophia

**Citing this paper**

Please note that where the full-text provided on Manchester Research Explorer is the Author Accepted Manuscript or Proof version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version.

**General rights**

Copyright and moral rights for the publications made accessible in the Research Explorer are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

**Takedown policy**

If you believe that this document breaches copyright please refer to the University of Manchester's Takedown Procedures [<http://man.ac.uk/04Y6Bo>] or contact [uml.scholarlycommunications@manchester.ac.uk](mailto:uml.scholarlycommunications@manchester.ac.uk) providing relevant details, so we can investigate your claim.



## On the Parfitian Thesis of Moral Responsibility

### Abstract

There seems to be a tension between determinism and moral responsibility such that, if determinism is true then perhaps we cannot be responsible for our actions. In his *On What Matters*, Derek Parfit tried to find a way to dissolve this tension by discussing a Kantian argument about the noumenal world. In recent years Parfit's argument has received some criticism, which has sought to undermine his argument while also making a variety of different claims about his actual views on this issue. In this paper, I argue that Parfit's argument requires modification; my proposals not only promise to make his argument clearer, but can also be used to answer some of his critics. I conclude by setting out a coherent argument—based on Parfit's writings—for defusing the tension between determinism and moral responsibility. We will see that the truth of determinism is not a threat for moral responsibility—at least in some senses.

**Keywords** Determinism, Responsibility, Suffer, Blame, Parfit

### 1. Introduction

Determinism is the thesis that all events and states of affairs are necessitated by antecedent causes—by how the world and laws of nature are. So, if we believe that our actions are events in time, it can be concluded that our actions are the inevitable consequences of causal forces which are beyond our control.

This fact seems to cause a problem for human responsibility. Historically, it was believed that, if we are responsible beings in a sense that we can be blamed or punished with respect to our wrong acts, we must possess a kind of control over our actions which enables us, at the moment of action, to act differently from what we actually did. But this seems to be in a direct conflict with determinism; because, according to the thesis, it would follow that given the state of mind of a person, it would have been impossible for her to have acted differently. In this way, it seems determinism makes it impossible for us to possess a kind of freedom which is required for moral responsibility.

Accordingly, it might be said that our acts cannot correctly be attributed to us in a sense that makes us their owner, and renders us responsible for our actions. These actions are determined by causal forces beyond our control; as Nagel puts it, for a person, “the effect of concentrating on the influence of what is not under his control is to make this responsible self seem to disappear, swallowed up by the order of mere events” (Nagel 2013 p.327). According to him, since “our acts must be viewed as part of the course of events,” the notion of responsibility and all related notions like punishment and blame become mere illusions (Nagel 2013 p.328).

In the second part of his monumental book *On What Matters*, Derek Parfit tries to show that with respect to the truth of determinism, human beings cannot be responsible for their actions in a sense that makes them deserve to suffer, but they can, indeed, be responsible in some other ways. His arguments have received criticism in recent years. In this paper I try to give an interpretation of his views to show that there is a Parfitian solution for the clash between moral responsibility and determinism that can withstand those criticisms, and which seems to be a promising way to deal with this problem.

In the first step, I briefly consider Parfit's argument for the thesis that no one deserves to suffer. Secondly, I mention some difficulties for this view—namely the criticisms of Zaibert and Smilansky. In the third section, I try to show that based on Parfit's ideas about personal identity, and a plausible distinction between two kinds of responsibility—responsibility in the sense that makes us deserve to suffer (hereafter RS) and responsibility in the sense that makes us deserve blame (hereafter RB)—we can devise a solution for the problem of responsibility. In this section I shed new light on Parfit's argument, while also adding an important modification to it which will help me respond to recent criticisms. I shall argue, then, that determinism could be compatible with moral responsibility in certain senses. Moreover, I will argue that accepting determinism might exclude RS, but that there is no contradiction between determinism and RB. I conclude that the truth of determinism gives us reason to be incompatibilists with respect to RS, but compatibilists with respect to RB.

## 2. Parfit's Argument for Moral Responsibility

There is a longstanding thesis that if we have responsibility, we must have responsibility in a sense that makes us deserve to suffer with respect to our wrong acts. Kant emphasizes this sort of responsibility and concludes that it shows that our actions cannot be events in time. Parfit summarizes Kant's argument as follows—what follows is a paraphrase (Parfit 2011 pp.258-264):

- A) If our acts were events in time, they would be determined by prior causes.
- B) If our acts were determined by prior causes, we could not have acted differently at the time of action.
- C) Ought implies can, so if we ought to have acted differently, this implies that we could have acted differently.
- D) Our acts could be wrong if we could be responsible for these acts in a sense that make us deserve to suffer.
- E) We could not have acted differently, so none of our acts could be wrong.

Therefore

- F) We do not have responsibility in a sense that makes us deserve to suffer and morality is an illusion.

However

- G) We have such responsibility and morality is not an illusion.

So

Our acts are not merely events in time.

Accordingly, drawing on the intuition that we are morally responsible agents, Kant concludes that our will is to be identified with our noumenal selves. Using responsibility in this sense has some merit; for example, it shows why children and the insane cannot act wrongly—Intuitively, it can be said that children and the insane do not deserve to suffer, no matter what they did; From this, we can conclude that they cannot be responsible in the sense that makes them deserve to suffer. Parfit, however, thinks that there could be another sense, according to which we can be morally responsible for our wrong actions even though we never deserve to suffer. In this way, he attacks premise D. Our ability to reflect and act on moral reasons, Parfit believes, is sufficient for holding us responsible.<sup>1</sup> After showing that for us to be moral agents we do not *need* to use the idea of responsibility in a sense that make us deserve to suffer, Parfit argues that we indeed *cannot* be responsible in this latter sense because no one can deserve to suffer. To demonstrate this claim, Parfit again uses another Kantian argument, showing that our acts are not merely events in time. The key feature of this argument is that if we were responsible for our acts in a sense that makes us deserve to suffer, we must also be responsible for our character. To be responsible in this sense, according to Kant, everyone must be able to freely choose her character—whether good or evil—since our actions depend on our character. Parfit restates this argument as follows—what follows is a paraphrase (Parfit 2011 pp. 267-269):

- A) We act in character—our actions depend on our character or the kind of person who we are.
- B) At any given time, to be responsible in a sense that makes us deserve to suffer, we must be responsible for the person who we are.
- C) If our acts were merely events in time, to be responsible for the person who we are, in any given time, we must be responsible for our past actions which made us the kind of person who we are—in that time.
- D) To be responsible for those actions, we must be responsible for the kind of person who we were at that time.

---

<sup>1</sup> And, of course, because of children's and the insane's inability to reflect and act on such reasons, they are not responsible in this sense and cannot act wrongly.

E) And to be responsible for the kind of person who we were at that time, we must be responsible for the actions which we did in further past.

F) This chain can be continued to the time when we were a child or even a fetus.

G) When we were a child we could not be responsible for our actions or the kind of person who we were.

So

J) We cannot be responsible for the person who we are, and we cannot be responsible for our actions in a sense that make us deserve to suffer.

But

R) We can be responsible for our actions in a sense that makes us deserve to suffer.

Therefore

S) Our acts are not merely events in time.

Parfit believed, as Kant himself declares, that the conclusion of this argument—namely noumenal freedom—is incomprehensible. We cannot expect ourselves to understand our freedom in a noumenal timeless world; all we can comprehend is our spatio-temporal world. Moreover, Parfit claims, this conclusion is unintelligible because one of its consequences is that in every trivial event which has ever occurred, the decisions of every human being who ever existed are implicated. Yet it is highly counter-intuitive that the decision of so many rational beings—some of them good and some of them bad—are what have brought about the whole sequence of the world just as it is now. Since the argument is valid and its conclusion is false, we need to reject one of premises. Parfit claims, we should reject either premise (J) or (R). According to him, it is intuitive to reject premise (R), and thus the argument would be altered as follows: (Parfit 2011 pp. 269-270).

“(J) if our acts were merely events in time, we could not deserve to suffer.

We can add

(T) Our acts *are* merely events in time.

Therefore

(U) We cannot deserve to suffer.”

In this section, then, I briefly considered Parfit’s argument. In the next two sections, other details of his thesis will be stated, criticized, and defended.

### 3. Criticisms

In this section I consider some important criticisms of Parfit’s thesis. Leo Zaibert, in his article “On the Matter of Suffering: Derek Parfit and the Possibility of Deserved Punishment” Zaibert 2014, tries to undermine Parfit’s thesis with four different arguments—two of which seem particularly important. Firstly, he claims that Parfit does not tell us what it is that makes RS different from other kinds of responsibility—other than that pain is bad. According to him, Parfit here is selectively picking this sense of responsibility because it serves his purpose well. Parfit is an incompatibilist with respect to RS and a compatibilist with respect to RB, but he does not give us a decisive reason for drawing such a distinction; and this makes his position unstable (Zaibert 2014 p.5).

Moreover, Zaibert claims, Parfit’s discussion about children and the insane gives us a reason to think he completely rejects desert, and this makes his view not only unstable but also contradictory:<sup>2</sup> “Unfortunately for Parfit, the facts that explain why children and (some of) the insane may indeed be incapable of deserving to suffer also explain why they are incapable of deserving anything else” (Zaibert 2014 p.5).

---

<sup>2</sup> Because Parfit believes we can deserve many things—except to suffer.

As mentioned, Zaibert claims that Parfit's only reason for claiming a difference between RS and RB is that suffering is bad; and because he only singles out RS—in the absence of any other reason for differentiating between RS and RB—it can be said that, according to Parfit, the badness of suffering is what makes it undeservable. With respect to this, Zaibert tries to show, even if we believe that pain is in itself bad, this would not always give us sufficient reason for avoiding pain; as a result, the badness of suffering cannot be used for showing that it is undeservable. In his words, the badness of suffering “need not be denied in order to insist that we can deserve to suffer” (Zaibert 2014 p.8). He appeals here to Moore's notion of *organic wholes*. According to Moore, there are good wholes that cannot exist unless some bad parts—namely suffering—exist too. In other words, the existence of suffering is a necessary condition for the existence of the good whole; and not just a means for it. The absence of the part would result in the destruction of the good whole—thus, the value of the part is not merely instrumental (Moore 1993).

So the existence of “good wholes” gives us reason in some cases to prefer states of affairs which contain more pain, because this is the only way of reaching those good wholes. For example, if a driver blamelessly kills a child, we expect him to be sad and feel remorse for the consequence of her action—although we do not blame or punish her. We would prefer such a world over another world which is exactly the same except that when a driver blamelessly kills a child, she does not feel remorse and as a result endures no pain. We prefer the former world, although we know that it contains more pain. We do this because this world has more intrinsic value<sup>3</sup>. Zaibert continues by remarking that if Parfit can admit that “it is good that the blameless driver should suffer (as a result of his regret), why should he not also admit that it is good that the blameworthy driver should suffer (as a result of his punishment)?” (Zaibert 2014 p.11).

However, in his “Parfit on Free Will, Desert, and the Fairness of Punishment,” Smilansky argues that Zaibert's claim that Parfit was an incompatibilist with respect to RS and a compatibilist with respect to RB is highly implausible. Parfit, he says, is clearly a hard determinist who cannot be a compatibilist in any sense. He puts forward two reasons for this: First, he argues, very briefly, that Parfit must be a hard determinist if he wants to effect a reconciliation between different theses in normative ethics. In his words “This would be threatened if it turns out that robust pre-institutional desert intuitions (namely, intuitions according to which the requirement for desert is basic and can precede the agreements formed by social institutions) can indeed be defended” (Smilansky 2016 p.140).

Smilansky's other reason is a textual one: “Parfit is clearly a hard determinist; rejecting free will-based moral responsibility and desert almost wherever they matter, and not only considering suffering” (Smilansky 2016 p.140). To show this he refers to some of Parfit's own utterances, like “no one could ever be responsible, I believe, in any way that could make them deserve to suffer. Nor, I believe, could anyone deserve to be less happy” (Parfit 2011 p.264); “When people treat us or others wrongly, we can justifiably be indignant...[However] We cannot justifiably have ill will towards these wrong-doers, wishing things to go badly for them” (Parfit 2011 p.272). If this is true and Parfit actually was a hard determinist, then Zaibert's first criticism would be undermined.<sup>4</sup>

In addition to this, Smilansky challenges Parfit's claim that although no one can deserve to suffer, punishment can be fair. If we believe—like hard determinists—that everyone is the person who she is because of heredity and the environment, and that her intentions, desires, etc. are outside her control, then punishing a criminal would be no fairer than punishing an innocent (Smilansky 2016).

In this section, I have considered several criticisms of Parfit's view, and it seems his thesis faces some critical issues. In the next section, I will try—based on Parfit's works—to articulate an argument for reducing the tension between determinism and responsibility which can also answer those criticisms.

#### **4. A Parfitian Thesis about Responsibility**

As we saw, a problem for moral responsibility is that it seems that, in the absence of a libertarian free will, our acts cannot be attributed to us in a sense that makes us responsible for them. Therefore, if I want to devise a Parfitian thesis about responsibility, I must show that our actions can in fact be attributed to us in a sense that

---

<sup>3</sup> Because there is intrinsic value in this kind of feeling remorse.

<sup>4</sup> As we shall see, this is not true.

makes us the owner of them. To show this we must be able to determine when it is correct to say that a person did a certain action. So, in turn, we must determine what sort of thing a person is (e.g. a body, an animal, etc.)—because in this way we will be able to say when a *person* did a certain act, as opposed to causal forces beyond her control. In reaching this goal, I appeal to the notion of a person which appears in personal identity (PI) debates. In this realm, philosophers try to define what it is that makes person X at time T<sub>1</sub>, one and the same person as person Y at time T<sub>2</sub>. If we could find the criterion of PI, then we could say what a person is.

A widespread thesis is that what is important for PI is a psychological criterion—which is also Parfit's view.<sup>5</sup> In this view, person X has PI with Y, if X is psychologically continuous with Y. Psychological continuity is an interrelated chain of psychological connectednesses, where psychological connectedness is a sufficient number of direct and strong psychological relations—like memories, intentions, beliefs, character, etc. (Parfit 1987). Now, if we can indeed say that it is psychological continuity that makes X and Y at different times one and the same person, we can conclude that a person is comprised by her psychological aspects—for simplicity, her character.

At this point we can argue that when an act is produced by a person's character—and not by the pressure of foreign forces like brain-washing—this act can be attributed to this person in a sense that makes her the owner of that act, and responsible for it. This is a perfect match with our intuitions. For example, consider a case in which a scientist examines two persons who are exactly the same—say a person and his duplicate. Suppose our scientist dramatically changes the character of one of them, but he leaves the other unchanged. Now, if these two persons do a certain wrong act, we believe intuitively that while the first person is not responsible for his action, the second, the one that is unchanged, is morally responsible.

In *What We Owe to Each Other*, Thomas Scanlon introduces a similar view. In his opinion, the simple fact that causal forces act through us cannot be a threat for our responsibility. In his words:

The reasons I have listed for preferring principles that make what happens to us depend on the ways we respond when presented with alternatives are not undermined if it turns out that these responses have causes outside us. As long as these causes affect our responses only by affecting what we are like, it will remain true that these responses can be good predictors of what will bring us enjoyment or advance our aims. ... this fact does not make them less a part of us and hence does not diminish the value of choices that express them. (Scanlon 2000 p.255)

We can claim, then, that unless causal forces affect our actions by affecting *our character*—i.e., by exerting an indirect influence—this will not directly cause a problem for our responsibility. We can be responsible, since the fact that causal forces influence our character, do not make our actions which stem from that character less a part of us. In this way, we can be responsible for our actions<sup>6</sup>.

A person's character is not something different from her, and when she acts based on her character this act can be correctly attributed to her in a sense that makes her the owner of this action and responsible for it. To illuminate this, consider the following example. Throughout their lives, human beings' bodies grow by eating different kinds of food. These foods, before being eaten, have a separate existence from the body—they can be pointed to and noticed. But when a person eats these foods and absorbs them, they become a part of her body and no longer have a separate existence. You cannot, for example, point to a part of your body and say this is a salad which I ate last year. That salad is gone, turned into a part of your body—your physical structure.

The same can be said about causal forces and our psychological structure. When these causal forces shape our character—our intentions, hopes, beliefs, etc.—this psychological structure no longer has a separate existence. Consequently, when a person acts in character, we cannot say that it was causal forces and not the person who did it. These acts are done by her and she is responsible for them.

---

<sup>5</sup> Although Parfit claims that personal identity is not important, whereas psychological continuity and connectedness are, this is not important for our purpose.

<sup>6</sup> Still, I need to answer the question that: in what sense, exactly, we can be responsible?

At this point another problem arises, namely the problem of self-creation<sup>7</sup>:

(A) If a person must be responsible for her actions which are based on her character, firstly, she must be responsible for her character.

(B) In order to be responsible for one's character, one must be able to freely choose one's character and create herself.

(C) No one can create herself in this sense

Therefore

No one can be responsible for her actions.

To meet this problem, I must show that responsibility does not depend on self-creation, or at least that there is a sense of responsibility that can survive this problem. As I mentioned, Parfit makes a distinction between two kinds of responsibility—in a sense that makes us deserve to suffer (i.e. RS), and in another sense (i.e. RB)—and believed that while RB can be defended, RS cannot, because no one deserves to suffer. So, at this point, I have to explain what it is that makes suffering or punishment undeservable. Is there any real difference between RB and RS, or this is just an ad hoc modification? Unfortunately, Parfit does not give us a clear reason for this distinction; nevertheless, it is plausible that there is a real difference and that it is related to the presuppositions of these two senses of responsibility.

Moral blame, in the sense that I use it here, is just a kind of criticism. Morality tells us how the world should be and moral blame is nothing more than a part of this. Moral blame tells a person how she should be, why a particular action or personal trait is not right, and how they must be modified. Moral blame should not be stated with rage and indignation, because in this manner it already contains a kind of punishment. Now, if we understand moral blame along these lines, the most important presupposition would be that the person is able to change in the future, with the help of moral blame and her reflection on it. This is completely compatible with determinism, because determinism does not reject the possibility of change in the future—otherwise it would reduce to fatalism. In this sense of responsibility, self-creation and its related notions are completely irrelevant. As Scanlon puts it:

If the Causal Thesis is true, and we are “like computers” in the very general sense that our mental lives depend on underlying causal processes, it will nonetheless remain true that we can communicate with each other in moral terms and that our behavior will be influenced by this kind of dialogue in just the way that it is now. (Scanlon 2000 pp.281-282)

For example, consider a case in which a 24-year-old woman, who always does her best to live a healthy life—eats healthy food, exercises regularly etc.—has a fatal genetic heart disease which past check-ups by different physicians has failed to diagnose. Moreover, suppose that if she does not get surgery before turning 25, she will die. Now, it is clear in this case that telling her about the situation is not related to the fact that her heart disease was beyond her control and or that she did not have a lifestyle which would have led to such a situation—considerations like this are completely beside the point. The only important fact here is that this woman can be cured in the future.

The possibility of change in the future is the only thing needed for RB. Parfit seems to have had such a belief in mind when he talked about the freedom which morality requires.

What we do often depends on our beliefs about what we ought to do. And if we come to believe that some act of ours was wrong, or irrational, because we ought to have acted differently, this belief may lead us to try to change ourselves, or our situation, so that we do not act wrongly, or irrationally, in this kind of way again. These changes in us or our situation may affect what we later do. . . . If we come to believe that we ought to have acted differently, this change in our beliefs may cause it to be true that in similar cases, without any miracle, we do in the future act differently. That is enough to make it worth asking whether we ought to have acted differently. (Parfit 2011 p.260)

---

<sup>7</sup> See Strawson, 2013, pp.313-317.

Although we can conclude that determinism is not a threat for RB, things are somewhat different with respect to RS. Alongside the presupposition which is stated in RB, there is another presupposition in RS that is not compatible with determinism. In punishment we usually pursue two purposes: first, to show the person as well as others that her action is wrong—this is our aim in RB—and second, to ensure that the person who did the wrong act pays for her action.<sup>8</sup> As I said, the first purpose and its presupposition are compatible with determinism; so, if there is an inconsistency with determinism, it must be related to the second. Here, our presupposition is that the person who acts wrongly could have acted differently, even with the same state of mind and character, at the time of action—this is the categorical sense of “could”. From this presupposition we can conclude that, because she did not act differently, she must pay for her action and therefore be punished.

But, categorical sense of could is not consistent with determinism, according to which all events and states of affairs are necessitated by antecedent causes. Therefore, this presupposition, which contains categorical sense of “could”, is in direct contrast with determinism. Given a woman’s state of mind at any given time in the past, it would be impossible for her to have acted differently. She could have acted differently only if her state of mind had been different—this is the hypothetical sense of “could”<sup>9</sup>. Now, if we admit that one of RS’s presuppositions is incompatible with determinism—the one which contains categorical sense of could—and also suppose that determinism is true, then we have to conclude that no one can be responsible in a sense that makes her deserve punishment or to suffer.

I believe that this reason for a differentiation between RB and RS was the missing part of Parfit’s argument. Suitably supplemented, we can see how his view can withstand all the criticisms mentioned in the previous section. Zaibert’s first criticism was that Parfit self-servingly differentiates between two kinds of responsibility, without giving a decisive reason for this. As we saw, there is a real difference between RB and RS, and the difference is related to their presuppositions. Moreover, Zaibert claimed that, contrary to Parfit, the badness of pain does not always give us a reason for avoiding it; therefore, he does not show that no one deserves to suffer. But, as I showed, to claim that no one deserves to suffer, Parfit does not need to refer to the badness of suffering—he only needs to refer to the presuppositions of punishment.

Moreover, it must be said that Parfit does not reject desert completely—only desert related to suffering—and so that Zaibert’s claim is false here. Parfit’s discussion about children and the insane, as mentioned in section 2, is not intended to show why people cannot deserve to suffer, but rather that for us to be responsible and be moral agents we do not need the notion of punishment or suffering: the ability to act and reflect on moral reasons is enough.<sup>10</sup> So, this cannot be used for showing that Parfit rejects desert completely. In the very same paragraph, Parfit points out that:

So we should reject Kant’s assumption that, for us to be moral agents, we must be responsible for our acts in some way that could make us deserve to suffer. We can coherently believe both that our acts can be right or wrong, and that no one could deserve to suffer. (Parfit 2011 p.264)

We do not say that a driver who blamelessly killed a child should suffer because she must pay for her action; and in this way, as Parfit mentions, we do not have ill-will toward her. We prefer a world in which blameless drivers suffer (as a consequence of their regret) because this shows our love and affection for each other, and can also make us more cautious and move us to prevent such situations from occurring in the future—needless also to mention that all these are instrumental values of such suffering and we do not say that people “deserve” it in such cases. So, this example cannot justify the thesis that suffering can be deserved. In the case of the blameworthy driver, we cannot want her to suffer punishment—paying for her action—because, given her character, it was impossible for her to act differently. We blame her in order to show how her character—which she did not choose freely—led to a disaster, and we ask her to change in future.

---

<sup>8</sup> Of course there are other purposes, but they are not relevant here.

<sup>9</sup> See Parfit 2011, p.260.

<sup>10</sup> This was of great importance for his aim to establish a difference between RS and RB, because if responsibility was dependent on punishment, we could not be responsible in any other senses.



There is another problem which is mentioned by Smilansky. Parfit, he claimed, could not be and indeed was not in any sense a compatibilist; as we saw in the quotation from Smilansky in section 2, Parfit believed in hard determinism and therefore was an incompatibilist. But, Smilansky does not give us a real argument for why Parfit cannot be a compatibilist, except that this would threaten his main aim of “the overall reconciliation of normative ethics” (Smilansky 2016 p.139). If you want to attribute such a striking claim to someone, it might seem that you owe him a real argument.

What about his textual argument? On the basis Parfit’s actual texts, can we claim he was an incompatibilist? Among all references which Smilansky makes to Parfit in support of his claim, the most interesting one is this: “no one could ever be responsible, I believe, in any way that could make them deserve to suffer. Nor, I believe, could anyone deserve to be less happy” (Smilansky 2016 p.140). This is interesting, because if you consider the whole paragraph, it would support just the opposite claim. Indeed, in this paragraph Parfit says:

According to premise (J), if our acts were merely events in time, we could not be responsible for our acts in this suffering-deserving way. This part of Kant’s view is, I believe, a profound truth. *We can be morally responsible in several other ways, or senses, but no one could ever be responsible, I believe, in any way that could make them deserve to suffer.* Nor, I believe, could anyone deserve to be less happy. (Parfit 2011 p.264; my emphasis)

Here, Parfit clearly differentiates several kinds of moral responsibility and only excludes responsibility in the sense that makes us deserve to suffer, as undeservable. Moreover, on the next page, Parfit claims:

The kind of freedom that morality requires is, I have claimed, *compatible* with determinism. . . . As Kant assumes, however, this kind of freedom is not enough to justify the belief that we can deserve to suffer for what we did. Kant *here* rightly rejects what we can call compatibilism about desert. (Parfit 2011 p.265; my emphasis)

Here again it is clear that with respect to the word “here” and the claim that morality is *compatible* with determinism, Parfit differentiates RB from RS and rejects only the latter. As Zaibert mentions, at the end of this chapter Parfit also claims that “We can deserve many things, such as gratitude, praise, and the kind of blame that is merely moral dispraise. But no one could ever deserve to suffer” (Parfit 2011 p.272). This distinction is present almost in every argument that Parfit proposes in this chapter; it would be highly implausible to say that Parfit did not believe in it, and that its presence is purely due to negligence. As a result, if we undermine Smilansky’s first claim—the claim that Parfit was actually an incompatibilist—we thereby undermine his other claim which presupposes it, namely the claim that, based on Parfit’s thesis, punishment cannot be fair.

But this may not be enough, because as we showed, Parfit actually is an incompatibilist with respect to punishment and in this case, we can accept Smilansky's claim about Parfit. However, contrary to Smilansky, I believe this is not a problem for, nor an inconsistency in Parfit's thesis. So, I have to show that punishment of wrongdoers, cannot be unfair, even though it is not deserved. In nearly all cases, there is a strong connection between fairness and deserve. It is fair to behave with someone in certain ways, or it is fair for something to happen to a person, when she deserves to be treated that way or deserves to be in that situation. But, there are cases in which such a connection does not exist. To show this, consider the following example which is taken from Scanlon<sup>11</sup>.

consider a woman who exposed to radioactive radiations, as a consequence of her behavior. Suppose the process in which the woman infected was a necessary process of transporting nuclear waste to a new site. Governments did all they could to protect people and informing them to stay far away from transporting area. The woman's curiosity led her to ignore warnings and go to the vicinity of transportation site where she got infected. In this way, although we can say that she is not justified in complaining about her infection saying it is unfair for her to be infected, we can also, plausibly, claim that it is not a good thing that she is infected as a result of her ignorance; it is not the case that she deserved that.

---

<sup>11</sup> See Scanlon, 2000, pp.263-267.

This is clearly a case in which something that is not deserved, fairly happens to a person. But why is that so? I believe there is two condition in cases like this which makes them different from normal cases. First, the action or the happening is inevitable—it is necessary. Second, the person who is treated in these ways or exposed to an event like this cannot complain—because we, the people who are relevant to the happening, did all we could to not behave with the person that way or to prevent the event from happening.

Another case which contains these conditions is quarantine<sup>12</sup>. When a person got infected by a contagious disease, it would be necessary to quarantine her, and if the governments did all they could for preventing that person from getting infected, she cannot complain and say that her quarantine is not fair. But we can plausibly say that she does not deserved that. When we consider punishment, we can see that the situation is the same. it seems without an institution of punishment, our societies and social life are bound to fail—punishment is a necessity of our social life. Moreover, in a society in which government and people did their share to prevent wrongdoers from doing criminal activities—say with providing them with a satisfying level of well-being and education—it would not be unfair to punish wrongdoers who do not deserve punishment. So, now we can claim that we showed how punishment of wrongdoers cannot be unfair, even though they do not deserve to be punished.

Zaibert and Smilansky, it seems, think Parfit's distinction between RS and RB is due to fault or negligence, and so try to undermine his argument. In Zaibert's words:

When he claims that Kant believed that “if our acts were merely events in time, we could not be responsible for our acts in this suffering-deserving way” (a view which Parfit considers a “profound truth”) (I, 264), Parfit is being understated to a fault. Kant's view is that if our acts were merely events in time, we could not be responsible for our acts in any deserving way. (Zaibert 2014 p.5)

But, as I showed, there is a real distinction between different senses of responsibility, and each criticism which presupposes this negligence can be answered. In the end, then, I can claim that I am able to present and defend a Parfitian thesis about moral responsibility. There are, clearly, other criticisms which I could not mention in this article; but still, defending Parfit's thesis against some challenging arguments is a sign of progress, and justifies us in thinking that we are on the right track.

## 5. Conclusion

In this article I considered some criticisms of Parfit's views about responsibility and suffering, and tried to show how, based on Parfit's thesis about personal identity and a distinction between RB and RS, we can answer these criticisms as well as overcoming a traditional problem for responsibility—namely the problem that, given the truth of determinism, our acts cannot be attributed to us in a sense that makes us responsible for them. My answer had two parts; first, I argued that provided our actions depend on our character and we act in a certain way because of our psychological structure, our actions can be attributed to us in a sense that makes us responsible for them.

Secondly, in addressing the self-creation problem, I argued that this kind of attribution is enough to make us responsible in a sense that makes us deserving of blame while not deserving to suffer. My reason for differentiating these two kinds of responsibility was a difference between their presuppositions—while RS supposes that a person could have acted differently in the past, RB does not have this supposition. We saw that this distinction can solve the problem of self-creation and also is able to undermine Zaibert's criticisms, namely that Parfit's distinction is arbitrary and that, contrary to Parfit, the badness of pain does not always give us a reason for avoiding it. I also showed, contrary to Smilansky, that in fact Parfit was a compatibilist with respect to RB and an incompatibilist with respect to RS—based on his writings in *On What Matters*. Moreover, with the help of two examples, I undermined his other claim that because Parfit is a hard determinist, he cannot argue for the fairness of punishment.

In the end, I can claim that Parfit's view about responsibility is a coherent and strong thesis which not only is able to solve some crucial problems of morality, but can also withstand criticism. A merit of this view is that it uses the insights of both compatibilists and incompatibilists to save the sense of responsibility without presupposing “odd metaphysical notions that have long since been out of fashion and that are admittedly most difficult to comprehend

---

<sup>12</sup> See Pereboom, 2013, pp.49-78.

clearly” (Taylor 2013 p.310). In this respect, our final position is somewhat like Smilansky’s Fundamental Dualism; however, there are some important differences between them.<sup>13</sup>

### References

- Moore, G. E. (1993). *Principia Ethica*, ed. Thomas Baldwin. Cambridge: Cambridge University Press.
- Nagel, T. (2013). Moral Luck. In R. Shafer-Landau (Ed.), *Ethical theory: an anthology* (pp. 322-330). Chichester, West Sussex: Wiley-Blackwell.
- Parfit, D. (2011). *On What Matters* (Vol. 1). Oxford, UK: Oxford University Press.
- Parfit, D. (1987). *Reasons and Persons*. Oxford: Clarendon Press.
- Pereboom, D. (2013). Free Will Skepticism and Criminal Punishment. In T. Nadelhoffer, (Ed.), *The Future of Punishment*, (pp. 49-78). New York: Oxford University Press.
- Scanlon, T. M. (2000). *What We Owe to Each Other*. Boston, MA: Harvard University Press.
- Smilansky, S. (2000). *Free Will and Illusion*. Oxford: Oxford University Press.
- Smilansky, S. (2016). Parfit on Free Will, Desert, and the Fairness of Punishment. Oxford: Oxford University Press
- Strawson, G. (2013). The Impossibility of Moral Responsibility. In R. Shafer-Landau (Ed.), *Ethical theory: an anthology* (pp. 313-317). Chichester, West Sussex: Wiley-Blackwell.
- Taylor, R. (2013). Determinism and the Theory of Agency. In R. Shafer-Landau (Ed.), *Ethical theory: an anthology* (pp. 308-312). Chichester, West Sussex: Wiley-Blackwell.
- Zaibert, L. (2014). On the Matter of Suffering: Derek Parfit and the Possibility of Deserved Punishment. *Criminal Law and Philosophy*. doi:10.1007/s11572-014-9360-z.

---

<sup>13</sup> See Smilansky’s *Free Will and Illusion* (2000).