



**WILEY-  
BLACKWELL**



---

Moral Responsibility and the Metaphysics of Free Will: Reply to van Inwagen

Author(s): John Martin Fischer

Source: *The Philosophical Quarterly*, Vol. 48, No. 191 (Apr., 1998), pp. 215-220

Published by: [Blackwell Publishing](#) for [The Philosophical Quarterly](#)

Stable URL: <http://www.jstor.org/stable/2660294>

Accessed: 24/02/2011 13:15

---

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/action/showPublisher?publisherCode=black>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact [support@jstor.org](mailto:support@jstor.org).



*Blackwell Publishing* and *The Philosophical Quarterly* are collaborating with JSTOR to digitize, preserve and extend access to *The Philosophical Quarterly*.

<http://www.jstor.org>

## MORAL RESPONSIBILITY AND THE METAPHYSICS OF FREE WILL: REPLY TO VAN INWAGEN

BY JOHN MARTIN FISCHER

I greatly appreciate Peter van Inwagen's thoughtful and engaging critical study of my book *The Metaphysics of Free Will*.<sup>1</sup> In this brief reply I would like to sketch a strategy of response to his major criticism of my approach.

Peter van Inwagen and I agree that there is a very potent argument to the conclusion that causal determinism is incompatible with the sort of control that involves genuinely open alternative possibilities. But whereas van Inwagen believes that moral responsibility requires alternative possibilities (at some point in time), I do not; thus van Inwagen is an incompatibilist about causal determinism and moral responsibility, whereas I am not. I am a semi-compatibilist; I believe that causal determinism is compatible with moral responsibility, even if causal determinism rules out alternative possibilities.

I follow Frankfurt in contending that there is a set of examples which render it highly plausible that individuals can be morally responsible for their behaviour, even though they do not have (and never have had) genuinely open alternative possibilities. Van Inwagen correctly points out that my discussion of this sort of example, although detailed in some ways, is relatively schematic in other ways. He contends that my crucial claim about these examples – that they suggest that moral responsibility depends on the actual history of an action and not upon the existence or nature of alternative scenarios – is 'at best misleading' (p. 375). He believes that I have gone astray by being insufficiently precise about the Frankfurt-style examples.

More specifically, van Inwagen contends that there are 'relatively precise principles relating moral responsibility and the ability to do otherwise that are *not* refuted by Frankfurt-type examples' (p. 376). One such principle, according to him (*ibid.*), is as follows:

If it is a fact that *p*, an agent is morally responsible for the fact that *p* only if that agent was once able to act in such a way that it would not have been the case that *p*.

This principle (which I shall call the 'principle of possible prevention') implies that an individual with no alternative possibilities is not responsible for any fact. And van Inwagen further points out that if no one is morally responsible for any fact, then it

<sup>1</sup> P. van Inwagen, 'Fischer on Moral Responsibility', *The Philosophical Quarterly*, 47 (1997), pp. 373–81. The present paper is a slightly revised version of a reply I gave to van Inwagen at an 'author-meets-critics' session on the metaphysics of free will at the American Philosophical Association Central Division Meetings, Pittsburgh, Pennsylvania, April 1997.

would seem that our belief that there is moral responsibility at all, for anything whatever, would be illusory.

Similarly, he goes on to articulate the 'no-matter-what principle':

If it would have been the case that  $p$  no matter what free choices or decisions Alice had made, then it seems plausible to suppose that Alice could not be morally responsible for the fact that  $p$ .

Van Inwagen contends (p. 378) that this principle 'is extremely attractive, and ... Frankfurt-style examples do nothing to lessen its attractiveness'.

I concede that it is very difficult to construct knock-down arguments about such issues, but I disagree with van Inwagen about his principles relating moral responsibility for 'facts' with alternative possibilities – the principle of possible prevention and the no-matter-what principle. I shall now lay out some of the reasons why I take a different view from his about the relevant sorts of cases.

I admit that there are *some* cases – the sort typically invoked by van Inwagen – in which it is plausible to say that an agent is not morally responsible for a fact that he cannot prevent from obtaining (and never could prevent from obtaining). It is even very tempting in these cases to suppose that it is precisely that the agent cannot prevent the fact from obtaining that makes it true that he is not morally responsible for the fact. But I believe we ought to resist this temptation, because there are *other* cases in which it is plausible and natural to think that an agent *is* morally responsible for a fact, even though he never could have prevented the fact from obtaining.

Here is an example, 'Joint Assassins'. Sam and Jack each want to assassinate the mayor of their town. Each has his own nefarious and no doubt misguided reasons for wanting the mayor dead. They act entirely independently of each other – they do not even know about the other's existence. Each of them deliberates and acts in a way which apparently makes him morally responsible for his actions – neither is compelled, coerced, deceived, crazy, manipulated electronically, hypnotized, and so forth. Both Sam and Jack go to a city council meeting at the town hall, and simultaneously pull the triggers of their guns. Their bullets strike the mayor at the same time, and the processes leading from each bullet's hitting him to a sequence of life-threatening biological events are similar. Each bullet's hitting the mayor in the way it does is sufficient for the mayor's being killed by it. Moreover, the situation is such that neither individual could prevent the other from shooting and killing the mayor – perhaps each is wearing a bullet-proof vest and other protective equipment. We could add further specifications which would make it the case that neither individual *ever* had the opportunity to prevent the other from being in a position to shoot and kill the mayor, though the details would only clutter the discussion here.

Now it seems very natural and plausible for me to say that Sam is morally responsible for the facts that the mayor is shot and that the mayor is killed, even though he cannot (and never could) prevent these facts from obtaining. Given that their situations are symmetrical, it seems to me that Jack is also morally responsible for the facts that the mayor is shot and that the mayor is killed, even though he cannot (and never could) prevent these facts from obtaining. Thus although each is *fully* responsible for the facts, neither is *solely* responsible for them.

In the grip of such principles as the principle of possible prevention and the no-matter-what principle, one might say that indeed Sam is *not* morally responsible for the fact that the mayor is shot (and the fact that the mayor is killed). Of course, on this view Sam might still be morally responsible for his act of shooting and killing the mayor, and for the particular way in which the mayor is shot (and dies), and so forth. But not for the fact that the mayor is shot and that the mayor is killed: after all, Sam could not have prevented these facts from obtaining.

The problem is that, given the symmetrical situations of Sam and Jack, one would need to say the same things about Jack. So Jack too would not be deemed morally responsible for the facts that the mayor is shot and that the mayor is killed. Thus although the mayor is in fact shot and killed, no one would be morally responsible for the facts that the mayor is shot and that the mayor is killed. But this seems highly implausible to me.

So I think that it is most attractive to say, in a case of simultaneous overdetermination such as 'Joint Assassins', that Sam is morally responsible for the relevant facts, even though he could not have prevented them from obtaining. Further, I believe that if one does in fact say this about a case of simultaneous overdetermination, it is also plausible to say this about a similar sort of case involving pre-emptive overdetermination. (The Frankfurt-type cases are, of course, cases involving a certain kind of pre-emptive overdetermination.)

So now here is another example, 'Assassin'.<sup>2</sup> Granted, the case is a 'science fiction' scenario with fantastic elements. Reasonable persons could question the intelligibility of such cases, but it is a standard Frankfurt-type example the intelligibility of which is not at issue in the debate between van Inwagen and me. (A full defence of its intelligibility is beyond the scope of this paper.) Here, in contrast to the situation in 'Joint Assassins', Sam confides in his friend, Jack. Sam tells Jack of his plan to murder the mayor of the town in which they live. Sam is disturbed about the mayor's liberal policies, especially his progressive taxation scheme. Whereas Sam's reasons for proposing to kill the mayor are bad ones, they are *his* reasons: he has not been hypnotized, brainwashed, duped, coerced, and so forth. Sam has deliberated coolly, and he has settled on his murderous course of action.

Sam is bad, and Jack is no better. Jack is pleased with Sam's plan, but Jack is a rather anxious person. Because Jack worries that Sam might waver, Jack has secretly installed a device in Sam's brain which allows him to monitor all of Sam's brain activity and to intervene in it, if he desires. The device can be employed by Jack to ensure that Sam decides to kill the mayor and that he acts on this decision; the device works by electronic stimulation of the brain. Let us imagine that Jack is absolutely committed to activating the device to ensure that Sam kills the mayor, should Sam show any sign of not carrying out his original plan. Also, we can imagine that Sam can do nothing to prevent the device from being fully effective, if Jack employs it to cause Sam to kill the mayor.

Sam and Jack both go to a meeting at the town hall, and Sam methodically carries out his plan to kill the mayor. He does not waver in any way, and he shoots

<sup>2</sup> Introduced in J.M. Fischer and M. Ravizza, 'Responsibility and Inevitability', *Ethics*, 101 (1991), pp. 258–78.

the mayor as a result of his original deliberations. Jack thus plays absolutely no role in Sam's decision and action; the electronic device monitors Sam's brain activity, but it does not have any causal influence on what actually happens. Sam acts exactly as he would have acted had no device been implanted in his brain.

Evidently Sam is morally responsible for what he has done. Indeed, he is blameworthy for deciding to shoot the mayor, for shooting the mayor and for killing the mayor. Further, I find it extremely plausible and natural to think that Sam is morally responsible for the fact that the mayor is shot (and the fact that the mayor is killed), even though (given the counterfactual intervener, Jack) Sam could not have prevented these facts from obtaining. (Of course, as above, we could add details that would make it the case that Sam never could have prevented these facts from obtaining.)

I argued above that in 'Joint Assassins' – a case of pre-emptive overdetermination – Sam is morally responsible for (say) the fact that the mayor is shot. (If one resists this conclusion, one must apparently say that no one is morally responsible for this fact!) But if one says that Sam is morally responsible in a case in which two independent causal sequences operate and lead to the mayor's being shot, one should certainly say that Sam is morally responsible in a case in which *only one* such causal sequence operates. So, although I do not know how to *establish* my view decisively here, I am pretty confident that Sam is morally responsible for the fact that the mayor is shot in 'Assassin'.

And if these are the right things to say about such cases as 'Joint Assassins' and 'Assassin', surely one should say a similar thing about van Inwagen's famous example (p. 376) of Gunnar, Ridley and Cosser:

Cosser wanted Gunnar to shoot and kill Ridley, which Gunnar seemed likely to do; he intended to, and he had the means and the opportunity. But if Gunnar had changed his mind about killing Ridley, Cosser would have manipulated Gunnar's brain in such a way as to have re-established his intention to shoot Ridley. In the event, Cosser's 'insurance policy' turned out not to have been necessary, for Gunnar did not change his mind, and shot and killed Ridley 'on schedule'. Cosser played no causal role whatever in the sequence of events that led up to the killing.

In so far as the case is structurally parallel in relevant respects to 'Assassin', it is natural for me to say that Gunnar is indeed morally responsible for the fact that Ridley's children are orphans, even though he could not have prevented this fact from obtaining.

I have conceded that I do not have a knock-down argument for what I am inclined to say about these cases. I have tried to support my position by pointing to implausible results apparently entailed by other views. I can now also offer a possible explanation for the attraction of such principles as the principle of possible prevention and the no-matter-what principle. I believe that one can distinguish a descriptive and a modalized version of facts such as the fact that the mayor is shot, the mayor is killed, and Ridley's children are orphans. The modalized version of the fact that the mayor is killed, for example, is the fact that, given the circumstances, the mayor had to be killed (one way or another), or perhaps the fact that, if the

mayor were not killed by Sam acting on his own, he would be killed by Sam acting as a result of the intervention of Jack's device. I agree that Sam is *not* morally responsible for this modalized fact. But nevertheless Sam is indeed morally responsible for the descriptive fact that the mayor is shot. This fact could be 'realized' or brought about in various different ways, so perhaps it is the fact that the mayor is killed one way or another. But there is still a distinction between this purely descriptive fact and the following modalized fact: the fact that the mayor had to be killed, because if Sam were not to do it, then Jack would.

The distinction between descriptive and modalized facts is a bit delicate. But maybe its delicacy adds to the cogency of my proposed explanation of the plausibility of such principles as the principle of possible prevention and the no-matter-what principle. It is perhaps easy to fail to notice the distinction between descriptive and modalized facts, and if one thinks in terms of the modalized facts, the principles become attractive. But there is indeed a distinction between the two kinds of facts, and the principles are unattractive as applied to descriptive facts. So Gunnar may well not be morally responsible for the fact that if Ridley's children had not become orphans in one way, then they would have become orphans in another; but Gunnar is nevertheless morally responsible for the fact that Ridley's children become orphans.

It is quite clear that there are *some* examples in which an agent is not morally responsible for a fact the obtaining of which it is impossible for the agent to prevent. But in my view van Inwagen tends to focus on a *proper subset* of such cases -- ones which make it very tempting to think that the impossibility of preventing the fact from obtaining is what rules out the agent's moral responsibility for the fact. One such example is van Inwagen's case (p. 378) of the plague-stricken village. Here the agent is supposed to take a serum up-river to a plague-stricken village, but instead gets drunk and misses the boat. Taking the boat is the only possible way to get to the village. But soon after the boat leaves the dock it strikes a rock and sinks. Hundreds of villagers who would have been saved by the serum die.

Van Inwagen is confident that the agent is not morally responsible for the fact that hundreds of villagers do not get the serum and consequently die. And I agree. But there are cases, and there are cases. In a variation on van Inwagen's case, the agent gets drunk and misses the boat, but has a counterfactual intervener associated with him: if he were about to decide not to get drunk, he would be required to get drunk and miss the boat anyway. In this version the boat never strikes a rock, and arrives at the village without the agent. In this version, as in van Inwagen's original version, the agent cannot prevent its being the case that hundreds of villagers do not get the serum and consequently die. And yet in the revised version I am inclined to say that the agent *is* nevertheless morally responsible for this fact.

In my view, it is important to take a more nuanced view of the range of examples in which it is true that the relevant individual cannot prevent a certain fact from obtaining. In some such cases it is indeed true that the agent is not morally responsible for the fact in question, but in others I am inclined to think that the agent can fairly be held morally responsible for the relevant fact. In van Inwagen's plague-stricken village example, the agent is not morally responsible for the relevant fact;

but in the revised version of the example, as in such examples as 'Joint Assassins' and 'Assassin', the agent *is* morally responsible for the relevant fact (even though he cannot prevent it from obtaining).

This, at least, is what I am inclined to say. In a series of papers, as well as in my new book (with Mark Ravizza), I seek to provide an account of the difference between the two kinds of examples.<sup>3</sup> I cannot go into the details here, but I can at least sketch the leading ideas. In certain examples of an agent's bringing about the truth of a fact (or the obtaining of a state of affairs), one can distinguish two components: the mechanism that issues in a bodily movement, and the process that goes from that movement to some event in the external world. When the internal mechanism (leading to the bodily movement) is appropriately reasons-responsive, and when the process going from the pertinent bodily movement to the event in the external world is suitably sensitive to the bodily movement, then the entire (two-part) sequence is 'responsive'. When ascertaining the responsiveness to reasons of the actual-sequence mechanism leading to the bodily movement, one holds fixed the kind of mechanism that actually operates; similarly, in assessing the sensitivity of the actual-sequence process leading to the event in the 'external world', one holds fixed the kind of process that actually takes place. A responsive sequence contains two linked and interlocking sensitivities. My view, then, is that an agent can be morally responsible for the obtaining of a state of affairs in so far as that state of affairs issues from a responsive sequence. In van Inwagen's plague-stricken village example the process is not responsive; no matter how the agent moves his body, presumably the boat will sink and the villagers will die.

In contrast, in the revised version of the example the sequence *is* responsive. That is, we hold fixed the non-intervention of the counterfactual intervener, and we note that the agent's actual bodily movements may well be appropriately responsive to reasons. Further, holding fixed the actual sort of process (in the external world, as it were), we note that had the agent got on the boat (with the serum), the villagers would have been saved.

Now I recognize that I have here provided only the barest sketch of my approach to responsibility for consequences. My main contention in this brief piece is that there is an interesting difference between the two sorts of cases pertaining to moral responsibility for facts an agent cannot prevent – a difference that makes me think that an allegiance to such principles as the principle of possible prevention and the no-matter-what principle may come from (or at least be bolstered by) an unfortunate tendency to attend only to a proper subset of the relevant cases.

*University of California, Riverside*

<sup>3</sup> J.M. Fischer and M. Ravizza, 'Responsibility for Consequences', in Jules L. Coleman and Allen Buchanan (eds), *In Harm's Way: Essays in Honor of Joel Feinberg* (Cambridge UP, 1994), pp. 183–208; 'The Inevitable', *Australasian Journal of Philosophy*, 70 (1994), pp. 388–404; and *Responsibility and Control: a Theory of Moral Responsibility* (Cambridge UP, 1998), pp. 92–123.