

## Moral Responsibility, Freedom, and Alternate Possibilities

By: [MICHAEL J. ZIMMERMAN](#)

Zimmerman, Michael J. "Moral Responsibility, Freedom, and Alternate Possibilities," *Pacific Philosophical Quarterly*, 63 (1982): 243-254.

Made available courtesy of Wiley-Blackwell.

The definitive version is available at [www3.interscience.wiley.com](http://www3.interscience.wiley.com)

**\*\*\*Note: Figures may be missing from this format of the document**

### **Article:**

John Locke asks us to consider a case of the following sort.<sup>1</sup> A man—call him Peter—is carried, while fast asleep, into a room where there is another man—call him Paul—who is fast asleep too. Peter is locked in the room and will not, upon awakening, be able to leave unless and until somebody else unlocks the door. Eventually, Peter wakes up and is surprised but delighted to find himself in Paul's room—delighted not because he likes Paul, but because he dislikes him and is glad of the opportunity to annoy him. So, instead of attempting to leave, Peter stays where he is, patiently waiting for Paul to wake up, in the full knowledge that Paul will be very annoyed to find him there. And, soon, Paul does wake up and, as predicted, is enraged at Peter's presence. Peter finds this all richly rewarding.

I have, with philosophic license, embellished Locke's original case a little; but, so embellished, the case is instructive. Given the perhaps unnecessary assumption that Paul does not deserve to be treated by Peter in the manner just described, it might well seem that Peter is in large measure morally responsible for Paul's being annoyed. (Of course, whoever carried Peter into the room in the first place is presumably also morally responsible to some extent for Paul's annoyance; however, this is a separate issue that I shall not investigate here.) But it is also true that Paul's annoyance was, under the circumstances, inevitable. For Paul is always annoyed at seeing Peter, and Peter could not have left the room even if he had tried. So, in effect, Peter could not have done other than to annoy Paul.<sup>2</sup> But, if this is so, Locke's case, when thus embellished, apparently serves to show the following principle false:

(P1) A person is morally responsible for what he has done only if he could have done otherwise.

In a recent, influential, and controversial paper, Harry Frankfurt argues that P1 (which he calls the principle of alternate possibilities) is indeed false, and he bases his argument on cases essentially similar to the case just given.<sup>3</sup> He then goes on to make a very important point: P1 has played a dominant role in almost all recent (and, I would add, not-so-recent) inquiries into the so-called free-will problem; indeed, P1 is a principle to which both major parties to the debate—compatibilists and incompatibilists—have in the past assented, and it is the interpretation of P1 which appears to have been the main bone of contention between them. The implication, then, of P1's being false seems to be that the rug is pulled out from under the disputants' feet; and, if this is correct, Frankfurt's disproof of P1 proves very powerful.

Reaction to Frankfurt's paper has been varied. Some have claimed that Frankfurt has not succeeded in proving P1 false. Others have claimed that he has succeeded in doing so, but that the falsity of P1 does not undermine the debate between compatibilists and incompatibilists to the extent that Frankfurt apparently thinks that it has. My purpose here, however, is not to comment on these comments, but rather to try to clarify the issue by noting in some detail the main alternative positions concerning it which are available for adoption. For, given the complexity of the issue, such clarification seems to me desirable and has not yet, to my knowledge, been provided.

We should need no reminding—but in fact we do need it—that the dispute between compatibilists and incompatibilists is primarily concerned with the relation between freedom of human action and determinism. Why, then, has P1 played such a dominant role in this dispute, given that it mentions neither freedom nor determinism but appears to concern itself, instead, with the matter of moral responsibility? The answer is fourfold. First, the debate between compatibilists and incompatibilists is fueled by a common concern with accounting for the possibility of persons being morally responsible for certain of their actions (and the consequences of these actions). Second, both parties to the debate assume, roughly, that a person's being morally responsible for an action requires that he acted freely. Third, both parties assume, roughly, that, to act freely, a person must be able to do something other than that which he in fact does. And fourth, the compatibilist asserts, roughly, that a person can do something other than that which he in fact does even if determinism is true, while the incompatibilist asserts, roughly, that a person can do something other than that which he in fact does only if determinism is false. There are, then, apart from P1, three basic principles which appear to be operative in the dispute, and these may be put roughly as follows:

- (P2) A person is morally responsible for what he has done only if he did it freely;
- (P3) A person has freely done what he has done only if he could have done otherwise; and
- (P4) A person could have done otherwise than he did only if determinism is false.

And at this point we may make these observations: first, P1 may be regarded as a "contraction" of P2 and P3 (that is, P1 may be viewed as "short for" the conjunction of P2 and P3, in that it is entailed by them but drops all explicit mention of freedom of action, which acts as a sort of mediator between moral responsibility and the ability to do otherwise); second, both compatibilists and incompatibilists accept what is roughly captured by P2 and P3; third, P2, P3, and P4 jointly imply the following principle:

- (P5) A person is morally responsible for what he has done only if determinism is false;

and fourth, the incompatibilist accepts P5 because he accepts P4, while the compatibilist sees no reason to accept P5 because he rejects P4.

But to put matters this way is to put them too roughly. For P2 appears to be false; moreover, both compatibilists and incompatibilists, when careful about what they are saying, acknowledge that P2 appears to be false. Consider the following case; Peter freely takes a drug, which he knows from experience will induce in him an irresistible desire to commit violence, and then, as a result, runs wild and causes considerable damage to his immediate environment. It seems very plausible to say that Peter is morally responsible for running wild (and for the damage he thereby causes), even though, given that he was at the time under the irresistible influence of the drug, he did not act freely in running wild. If this is correct, then P2 is indeed false. Now, it is conceivable that some would seek to defend P2 in the following manner.<sup>4</sup> They might acknowledge that Peter did not act freely in running wild, and they might acknowledge that he deserves to be made to make reparations for the damage he has caused, but they might still insist that he is not morally responsible for running wild. This line of reasoning strikes me as incorrect, although how it might be disputed is not clear. One way to try to dispute it is to claim that the following principle is true:

- (P6) A person deserves to be made to make reparations for damage he has caused only if he is morally responsible for causing the damage.

P6 seems to me to be true, but no doubt the defender of P2 would object to it. (Indeed, it must be acknowledged that the law does not seem to subscribe to P6 in all cases where reparations for damage are concerned.) If so, I do not know what more could be done to convince the defender of P2 of its falsity. Nevertheless, I think it will generally be acknowledged that the foregoing case of drug-induced violence does succeed in showing P2 to be false.

But, even if P2 is shown false by the case of drug-induced violence, it nevertheless suggests that a principle closely related to P2 is true. This principle may be roughly stated as follows: a person is morally responsible for what he has done only if either he did it freely or he did it as a result of something else he did freely. But at this point roughly stating a principle is no longer adequate, and so I offer this more precise formulation of it:

- (P2') For any person  $p$ , any times  $t_2$ , and  $t_3$ , (where  $t_2$  is not later than  $t_3$ ), and any action  $a$ :  
 $p$  is morally responsible at  $t_3$  for doing  $a$  at  $t_2$  only if either
- (i)  $p$  does  $a$  freely at  $t_2$ , or
  - (ii) for some action  $b$  and some time  $t_1$  not later than  $t_2$ ,
    - ( $\alpha$ )  $p$ 's doing  $a$  at  $t_2$  is a consequence of  $p$ 's doing  $b$  at  $t_1$ , and
    - ( $\beta$ )  $p$  does  $b$  freely at  $t_1$ .<sup>5</sup>

This transformation of P2 into P2' does not, of course, adversely affect the intuition underlying P1, but it does suggest that P1 itself requires a more careful formulation. For, not only is it true that Peter did not act freely when running wild, it is also true that *at the time* he could not have done otherwise. But it does not follow that he could not *at any time* have avoided running wild at the time at which he did run wild. The point is—and this has now become a commonplace in the literature—that "can" requires two time-indices, one index specifying the time of "can" and the other specifying the time of the action in question.<sup>6</sup> For example, let us suppose that Peter takes the drug at  $t_1$  and as a result runs wild at  $t_2$ . While it is true to say that Peter cannot at  $t_2$ , do other than run wild at  $t_2$ , it is also true to say that Peter can at  $t_1$ , do other than run wild at  $t_2$ . Indeed, defenders of P1 will insist that it is precisely because of this latter fact that Peter may be said to be morally responsible for running wild. Hence, P1 should presumably be modified as follows:

- (P1') For any person  $p$ , any times  $t_2$  and  $t_3$  (where  $t_2$  is not later than  $t_3$ ), and any action  $a$ :  
 $p$  is morally responsible at  $t_3$  for doing  $a$  at  $t_2$  only if, for some time  $t_1$  not later than  $t_2$ ,  $p$  can at  $t_1$  do other than  $a$  at  $t_2$ .

Notice that it is of course true that, if  $p$  can at  $t_2$  do other than  $a$  at  $t_2$ , then there is a time  $t_1$  not later than  $t_2$  such that  $p$  can at  $t_1$  do other than  $a$  at  $t_2$ . In fact, it seems that the traditional debate between compatibilists and incompatibilists has been based partly on the shared premise that freedom of action requires that the person be able at the time in question to do other than that which he in fact does. That is, P3 should presumably be modified as follows:

- (P3') For any person  $p$ , any time  $t_1$ , and any action  $a$ :  
 $p$  does  $a$  freely at  $t_1$  only if  $p$  can at  $t_1$  do other than  $a$  at  $t_1$ .<sup>7</sup>

Finally, P4 and P5 also need, in the present context, to be made more precise, and their modification should presumably be as follows:

- (P4') For any person  $p$ , any time  $t_2$ , and any action  $a$ :  
if  $p$  does  $a$  at  $t_2$  and determinism is true, then there is no time  $t_1$  (not later than  $t_2$ ) such that  $p$  can at  $t_1$  do other than  $a$  at  $t_2$ .
- (P5') For any person  $p$ , any time  $t_2$ , and any action  $a$ :  
if  $p$  does  $a$  at  $t_2$  and determinism is true, then there is no time  $t_3$  (not earlier than  $t_2$ ) such that  $p$  is morally responsible at  $t_3$ , for doing  $a$  at  $t_2$ .

Unfortunately, even these somewhat complex reformulations are not adequate to the task at hand. Two key points were noted earlier about the unrefined principles P1 through P5: first, P2 and P3 jointly imply P1 and, indeed, P1 can be seen as a contraction of P2 and P3; secondly, P2, P3, and P4 jointly imply P5. But, because of the (necessary) introduction of the double time-index in P1' (and the corresponding complications concerning times in the other principles), this point no longer holds on translation from the unrefined principles to their reformulations. That is, although it is true that P2', P3', and P4' jointly imply P5', it is not true that P2' and P3'

jointly imply P1'. But this situation may be rectified by the provision of the following principle, which serves to "link up"  $t_1$  and  $t_2$  (when these are distinct) in the modified principles:

(P7) For any person  $p$ , any time  $t_1$ , and any action  $b$ :

$p$  does  $b$  freely at  $t_1$  only if, for all times  $t_2$  not earlier than  $t_1$  and all actions  $a$ , if  $p$ 's doing  $a$  at  $t_2$  is a consequence of  $p$ 's doing  $b$  at  $t_1$ , then  $p$  can at  $t_1$  do other than  $a$  at  $t_2$ .

I say "rectified," because the following point is true: P2', P3', and P7 jointly imply P1', and P1' may be seen as a contraction of these principles.<sup>8</sup>

Now, if P1 through P5 are reformulated in this manner and also supplemented by P7 (as it seems they should be, if justice is to be done to the complexity of the debate between compatibilists and incompatibilists), an interesting point emerges, and that is that the case of Peter, Paul, and the locked door turns out clearly not to be a counterexample to P1'. For, although it is true that, when he is in Paul's room, Peter cannot leave, we are not told that there is no time at all at which Peter can leave Paul's room.<sup>9</sup> The point is that it may be true, for all we are told, that Peter could have avoided becoming locked in Paul's room in the first place; in fact, it would be very odd if this were not the case. For instance, we may safely assume that Peter could have left town the previous day, in which case he would not have been a candidate for clandestine transportation to Paul's room; and, of course, if Peter had not been transported to and locked in Paul's room, he would not then have annoyed Paul. So, although it is true to say that, given that Peter annoyed Paul at  $t_2$ , he could not at  $t_2$  have done otherwise at  $t_2$ , it is false to say that there was no time  $t_1$  such that Peter could at  $t_1$  have done other than to annoy Paul at  $t_2$ ; and it is for this reason that, even if we grant that Peter is morally responsible for annoying Paul at  $t_2$ , the case, as it has been set up, does not serve to show P1' false.<sup>10</sup>

But the case can be modified so that it constitutes an apparently genuine counterexample to P1'. One such modification runs as follows. Peter and Paul are twins placed in a room at birth from which they cannot escape. They grow up unaware of their being forcibly confined. After some years, the twins become "responsible" human beings—that is, persons to whom it is appropriate to ascribe moral responsibility for certain of their actions (and the consequences thereof). In addition, Peter's very presence in the room has started to get on Paul's nerves, and Peter is aware of this; but, instead of deciding to leave the room (which is, it seems, something that he cannot and never could do, but something he nevertheless believes he can do), Peter decides at  $t_2$  to remain where he is, his express purpose being to annoy Paul thereby. Now it seems that at no time  $t_1$  could Peter do other than to annoy Paul at  $t_2$ ; and yet it also seems that Peter is in large measure morally responsible for Paul's being annoyed. If this is correct, then P1' is false.

The question now to be addressed is: what effect does this modified case of Peter, Paul, and the locked door have on the debate between compatibilists and incompatibilists, given that this debate seems to have been premised on the truth of P1'? Earlier, when presented in rough form, the debate was seen to rest on an agreement on the truth of P1 and to turn on a disagreement on the truth of P5 due to a disagreement on the truth of P4. This pattern remains on the more precise formulation of the debate. That is, on this formulation, the debate rests on an agreement on the truth of P1' and turns on a disagreement on the truth of P5' due to a disagreement on the truth of P4'. So, if P1' is false, the debate does indeed appear to be undercut. For even if P4' is true, it is not, in the absence of P1', sufficient to yield P5'; and so the common concern with accounting for the possibility of persons being morally responsible for certain of their actions (and the consequences thereof), which originally fueled the debate, would appear irrelevant to the debate when P1' is no longer a factor in it. But is P1' false, and, if it is, to what extent is the debate undercut? At this point I shall not attempt to provide a definitive answer to these questions; rather, I shall present what I take to be the main alternative positions open to anyone seeking to provide such an answer.

First, it is of course possible that someone should reject the modified case of Peter, Paul, and the locked door as a counterexample to P1'. Such a rejection could be based on one or the other of two claims. It might be claimed (i) that Peter *could* at some time  $t_1$  have avoided annoying Paul at  $t_2$  (or, at least, that it has not been *shown* that he

could not have avoided this), or it might be claimed (ii) that Peter is *not* morally responsible at any time  $t_3$  for annoying Paul at  $t_2$ . If either of these claims is correct, P1' remains intact, and the debate between compatibilists and incompatibilists remains unaffected.

The rationale for claim (ii) might be that Peter's willingly remaining in the room to annoy Paul does, under the circumstances, make him morally responsible for annoying Paul, but, had he not willingly done so, the circumstances would have been different; and the fact that, under these other circumstances, he would still have annoyed Paul does not entail that, under the actual circumstances, he could not have avoided annoying Paul.<sup>11</sup> Now perhaps, in the absence of an explication of the phrase "could have done otherwise"—and its analysis is notoriously controversial—this line of reasoning is defensible, but I have my doubts; for it certainly does seem to me that, whether one subscribes to a compatibilist or incompatibilist analysis of that phrase, it is pretty clear that, in the modified case, there is no time  $t_1$  at which Peter can do other than annoy Paul at  $t_2$ . But I shall not pursue this here.

The rationale for claim (ii) might be that, while it is indeed correct to condemn Peter's motives in annoying Paul (or, perhaps, to condemn him for these motives), it is nevertheless incorrect to ascribe moral responsibility to him for actually annoying Paul, and this is so simply because there was no time at which he could have avoided doing this (that is, because P1' is true).<sup>12</sup> This rationale is, I think, an interesting one, and it might be bolstered by the following observation. The modified case is set up so that Peter is unaware that his confinement is forced upon him. Why should this make a difference as to whether he is morally responsible for annoying Paul? The implication seems to be that, had he known of his confinement to the room—or even just believed that he was so confined—he would not have been morally responsible for annoying Paul, even though he desired to do so. But if the presence of a desire to annoy Paul is not, under the circumstances, sufficient to make Peter morally responsible for annoying Paul, why should such a desire conjoined with a (false) belief that one can do otherwise be so sufficient? This, I think, is an important question, but again it is not one that I shall try to answer here. It should be noted, however, that commentators on Frankfurt's paper appear generally to have overlooked such a rationale for claim (ii).<sup>13</sup>

If either of claims (i) and (ii) is correct, then P1' is safe and the debate is not undercut by the modified case of Peter, Paul, and the locked door. But if this case does succeed in showing P1' false, then the debate *is* undercut. But to what extent? This depends in part on just *what* is shown false by showing P1' false. It should be recalled that P1' may be viewed as a contraction of P2', P3', and P7, for these principles jointly entail it and may be seen to be "embodied" in it. So the question arises: given that the case shows P1' false, which of P2', P3', and P7 does it show false? There seem to me to be three main alternatives here. It might be claimed (iii) that the case shows P1' false but does not show any of P2', P3', and P7 false (although it remains true, of course, that if P1' is false, then at least one of P2', P3', and P7 *is* false); or it might be claimed that the case, by showing P1' false, shows P2' false; or it might be claimed that the case, by showing P1' false, shows P3' and P7 false. Of course, there are other combinations of P2', P3', and P7 which might be called on in this respect, but claims (iii), (iv), and (v) appear to incorporate the most obvious candidates.

I am not sure what the rationale for claim (iii) might be. It seems to me an odd claim, for it seems to pay no heed to the way in which the traditional dispute between compatibilists and incompatibilists has been set up; that is, it ignores the fact that the reason why P1' has been agreed to by all parties to the debate in the past is, or seems to be, because of a common agreement to P2', P3', and P7 (or something close to them). Nevertheless, ignoring the fact that P1' is a contraction of these principles has become a common practice. Certain commentators on Frankfurt's paper, for instance, have accepted his rejection of P1' but have sought to show that the debate between compatibilists and incompatibilists lives on in one form or another.<sup>14</sup> How to assess their remarks in this regard is, however, far from clear simply *because* they ignore the fact that P1' traditionally derives its status as a shared premise between compatibilists and incompatibilists *in virtue of* its being a contraction of P2', P3', and P7. For this reason, I shall not comment further on claim (iii) or on the commentators on Frankfurt's paper who, by their silence with respect to P2', P3', and P7, appear to subscribe to it.

The rationale for claim (iv) might simply be based on an insistence on the truth of P3' and P7. It might be said, roughly, that there just cannot be freedom of action in the absence of genuine alternatives, although there may of course be the illusion of such freedom if one is unaware of the absence of such alternatives. This contention has, in fact, been made many times, in one form or another, throughout much of the history of philosophers' attention to the free-will problem. If this contention is correct, and P1' is also false, then it follows that P2' is false. Now, if P2' is false, how does this affect the debate between compatibilists and incompatibilists? It seems to me that, in this case, the debate would be pretty effectively undercut. For, although the issue of whether freedom of action is compatible with the truth of determinism would remain unresolved, the common concern that originally served to fuel this debate would no longer be relevant to it. That is, however the debate were to be resolved, if P2' is false its resolution would appear to have no implication with respect to the truth of P5'.

Finally, claim (v) commands serious attention. It is, in fact, this claim to which I subscribe, although I shall not seek to defend it here; it is, also, a claim which, so far as I know, has been ignored by commentators on Frankfurt's paper; moreover, it is a claim to which both compatibilists and incompatibilists may subscribe without too much alteration of their original beliefs. The rationale for this claim might go as follows. P1' is indeed false, as the modified case concerning Peter, Paul, and the locked door shows. But this case does *not* show P2' false—for P2' is true—rather, it shows P3' and P7 false. That is, to put it roughly, the case shows that freedom of action does not require the ability to do other than that which one does.<sup>15</sup> What freedom of action requires, rather, is that the person who acts be, in some sense, the "source" of his actions, and one can be such a source even when one has no alternative but to act as one does. What is it to be the source of one's actions? It is here that compatibilists and incompatibilists will, indeed do, differ. Roughly, the compatibilist will say, as indeed he often does say, that freedom of action consists in acting in accordance with certain of one's own wishes or desires, perhaps only when these are coupled with a belief that one's options are relatively unrestricted.<sup>16</sup> (This rough characterization of the compatibilist view of freedom applies also to the more sophisticated hierarchical accounts of freedom that have recently appeared.<sup>17</sup>) The incompatibilist, on the other hand, will say, as indeed he always does say, that freedom of action requires lack of causal determination of one's action. If he is a libertarian to boot, the incompatibilist will locate the agent's being the source of his own free actions in the phenomenon of agent-causation, and this too might be coupled with a requirement that the agent believe that his options are relatively unrestricted;<sup>18</sup> if he is a determinist, the incompatibilist might simply say that there is no clear sense to be given to the notion of an agent's being the source of certain of his own actions. But, however these differences of opinion are to be worked out, P2' can still act as the shared premise in the debate. It is true that this debate is a modification of the traditional debate, in that P3' and P7 no longer feature as part of the premise shared by both parties; nevertheless the spirit of the original debate is preserved—its urgency is not at all diminished—for it is still fueled by a common concern with accounting for the possibility of persons being morally responsible for certain of their actions (and the consequences thereof). In fact, the debate is simplified by this modification. For, on this modification, the debate requires just P2', P5', and one other principle, namely:

(P8) If determinism is true, then no one ever acts freely.<sup>19</sup>

That is, given P2' as the premise shared by both parties to this modification of the original debate, the incompatibilist accepts P5' because he accepts P8, while the compatibilist sees no reason to accept P5' because he rejects P8. At this point, however, the question arises as *to why* the incompatibilist accepts P8, since he no longer accepts P3' and P7 and hence can no longer use P4' (which, I take it, he will still accept<sup>20</sup>) in his defense of P5'. That is, why does the incompatibilist regard determinism not only as sufficient to rule out anyone's ever doing other than what he does but *also* (and this is now an *independent* point) as sufficient to rule out anyone's ever acting freely? The answer can only be that he regards a person's being caused by antecedent events to do what he does as sufficient to rule out his being the source of his actions. This is perhaps a difficult position to hold, given the incompatibilist's rejection of P3' and P7, because—as the modified case of Peter, Paul, and the locked door shows—a person's being unable to do otherwise than he does is not regarded by the incompatibilist as being sufficient to rule out his being the source of his actions—at least this is so for any incompatibilist who thinks that it makes sense to say that someone can be the source of certain of his actions. And so any such incompatibilist who subscribes to claim (v) is committed to accepting the following principle:

- (P9) For any person  $p$ , any time  $t_1$ , and any action  $a$ :  
it is possible both that  $p$  cannot at  $t_1$  do other than  $a$  at  $t_1$  and that  $p$  is not caused by some antecedent event to do  $a$  at  $t_1$ .

Indeed, the modified case of Peter, Paul, and the locked door is, given claim (v), testimony to the truth of P9. Thus the incompatibilist who regards Peter as the source of his own action of annoying Paul is committed to saying that Peter was *not* caused by antecedent events to annoy Paul even though he could not do other than annoy Paul. Although this seems a tenable position to me, it is perhaps a difficult one to hold, and the incompatibilist's position is obviously open to some criticism on this score.

But I shall not pursue this line of inquiry further. Rather, let me summarize the main points made in this paper. The paper opened with a story about Peter, Paul, and a locked door. Certain rough morals were seen to be perhaps derivable from the story. But the morals were too rough and required refinement. When they were refined, it became clear that the original story required refinement also. Then, given the modified story, five main positions were seen to be available for adoption with respect to this story's effect on principle P1' and, through this, on the traditional debate between compatibilists and incompatibilists. Which position is the correct one was left open to question, although I noted my tendency to subscribe to that position captured in claim (v). The important point to note is that all of these five positions (and others too<sup>21</sup>) are available for adoption, for this seems to me to be a point that has been overlooked in recent discussion of the issue. Indeed, the issue is, I believe, clarified to a certain extent merely by noting this point.<sup>22</sup>

## NOTES

1. John Locke, *An Essay Concerning Human Understanding* (New York: Collier Books, 1965), pp. 149-150.
2. It might be said that Peter does not do anything when he annoys Paul, I disagree; I think Peter's remaining in the room is an action of his. However, if a more "positive" action is desirable (or, at least desired) for the sake of illustration, then the case may be modified so that, say, Peter punches Paul and thereby annoys him. Of course, the *inevitability* of his punching Paul would require more than just his being locked in the room; in fact, it would require something quite fantastic (such as his bruin's being monitored by an evil demon), and for this reason I shall stick with the more mundane example just given.  
It might also be said that, while Peter's remaining in the room is an action of his, had he tried to leave his remaining would then *not* have been an action. There is some plausibility to this, but again the case could be modified so that Peter is seen to have no alternative but to act in a certain way. (For a case of this sort, see David Blumenfeld, "The Principle of Alternate Possibilities," *The Journal of Philosophy*, 68 (1971), p. 341, n. 3.)
3. Harry G. Frankfurt, "Alternate Possibilities and Moral Responsibility," *The Journal of Philosophy*, 66 (1969), 829-839.
4. This point was brought to my attention by Douglas Husak.
5. Note that, given the explicit mention of times for precision's sake in P2', the verbs in this principle are to be understood tenselessly, while this is not the case with the verbs in P2. This point holds analogously for the principles that follow.  
Note also that the time at which  $p$  is said to be morally responsible— $t_3$ —is specified as not being earlier than  $t_2$ . This is in accordance with the fact that the sort of moral responsibility in question is "backward-looking" (or, at least, not "forward-looking"), as opposed to the sort of responsibility that H. L. A. Hart calls "role-responsibility," which is "forward-looking" (or, at least, not "backward-looking"), (See H. L. A. Hart, *Punishment and Responsibility* (Oxford: Oxford University Press, 1968), p. 211ff.) Note finally that, in saying that  $p$  is morally responsible at  $t_3$  for doing  $a$  at  $T_2$ , I am of course not saying that it is only at  $t_3$  that  $p$  is so responsible.
6. For a relatively early example of the application of two time-indices to "can"-contexts, see Keith Lehrer and Richard Taylor, "Time, Truth, and Modalities," *Mind*, 74 (1965), 390-395.
7. Someone who subscribes to the so-called prolific theory of actions might say that it is (almost) inevitable that, whenever a person does something, he also does something else, and hence that P3' is of no help in

distinguishing free from unfree actions. But P3' could easily be amended so that, even on the prolific theory, it states something useful. The following amendment should do:

(P3'') For any person  $p$ , any time  $t_1$ , and any action  $a$ :

$p$  does  $a$  freely at  $t_1$  only if, (or some action  $b$ ,  $p$  can at  $t_1$  do  $b$  at  $t_1$  without doing  $a$  at  $t_1$ .)

8. The introduction of P7 may seem an unnecessary complication, For the earlier pattern of implication may be preserved by replacing P1' with:

(P1'') For any person  $p$ , any times  $t_2$  and  $t_3$ , (where  $t_2$  is not later than  $t_3$ ) and any action  $a$ :

$p$  is morally responsible at  $t_3$  for doing  $a$  at  $t_2$ , only if either

(i)  $p$  can at  $t_2$  do other than  $a$  at  $t_2$ , or

(ii) for some action  $b$  and some time  $t_1$ , not later than  $t_2$ ,

( $\alpha$ )  $p$ 's doing  $a$  at  $t_2$  is a consequence of  $p$ 's doing  $b$  at  $t_1$ , and

( $\beta$ )  $p$  can at  $t_1$  do other than  $b$  at  $t_1$ .

That is: P2' and P3' jointly imply P1''; and P2', P3', and P4' (as before) jointly imply P5'. But, while I agree that anyone who accepts P1' is likely to accept P1'' (indeed, P1' *implies* P1'' when the former is conjoined with the following apparently unexceptionable principle:

(I) For any person  $p$ , any times  $t_1$  and  $t_2$  (where  $t_1$  is not later than and any action  $a$ :

if (i) ( $\alpha$ )  $p$  can at  $t_1$  do other than  $a$  at  $t_2$ , and

( $\beta$ )  $p$  cannot at  $t_2$  do other than  $a$  at  $t_2$ ,

then (ii) for some action  $b$ ,

( $\alpha$ )  $p$ 's doing  $a$  at  $t_2$  is a consequence of  $p$ 's doing  $b$  at  $t_1$ , and

( $\beta$ )  $p$  can at  $t_1$  do other than  $b$  at  $t_1$ ),

P1'' is in fact too *weak* for present purposes. For the consequent of P1'' does *not* imply that  $p$  can, at *any* time, do other than  $a$ , and yet such implication is surely required in the present context. Of course, this situation could be rectified by the provision of the following principle:

(II) For any person  $p$ , any time  $t_1$ , and any action  $b$ :

$p$  can at  $t_1$  do other than  $b$  at  $t_1$  only if, for all times  $t_2$  not earlier than  $t_1$  and all actions  $a$ , if  $p$ 's doing  $a$  at  $t_2$  is a consequence of  $p$ 's doing  $b$  at  $t_1$ , then  $p$  can at  $t_1$  do other than  $a$  at  $t_2$ .

But II is very similar to P7, and so greater simplicity is not achieved by this approach.

9. It should be pointed out here that those cases which Frankfurt (*op. cit.*) considers to be counter examples to P1 apparently fail to be counterexamples to P1' just as the case of Peter, Paul, and the locked door does. For, although it seems plausible to say that, in Frankfurt's examples, an agent  $p$  "could not have done otherwise," it turns out on closer inspection that the sense in which this is true is apparently this:  $p$  cannot at  $t_2$  do other than action  $a$  at  $t_2$ , and not this;  $p$  cannot at any time  $t_1$  do other than  $a$  at  $t_2$ . Nevertheless, I think that examples can be drawn up which will put P1' to the test in the manner intended by Frankfurt; the modified case of Peter, Paul, and the locked door, given below, is one such example.
10. The question arises, however, why we would presumably be willing to absolve Peter of moral responsibility for Paul's annoyance had he tried to leave the room before Paul awoke and found himself unable to do so, even though it would still have been true that Peter could at some time  $t_1$  have avoided annoying Paul at  $t_2$ . The answer, I think, is to be found in the observation that, for all times  $t_1$  at which Peter could do other than to annoy Paul at  $t_2$ , his annoying Paul at  $t_2$  was not foreseeable by him at  $t_1$ ; but also, for all those times  $t_1$  at which his annoying Paul at  $t_2$  was foreseeable by him, he could not at  $t_1$  do other than to annoy Paul at  $t_2$ . Exactly why this would constitute an adequate excuse for Peter's behavior at  $t_2$  is, however, a question that I shall not address here.
11. Cf. Robert Audi, "Moral Responsibility, Freedom, and Compulsion," *American Philosophical Quarterly*, 11 (1974), pp, 4-5.
12. This point was suggested to me by Stephen Hudson.
13. But cf. Blumenfeld, *op. cit.*, where certain remarks on pp, 343-345 suggest that some such point is at issue.
14. The following principles have, for example, been discussed as possible replacements of P1' as common ground for the compatibilist-incompatibilist dispute:
- (III) A person is not morally responsible for what he has done if he did it because he could not have done otherwise;



(IV) A person is not morally responsible for what he has done if he did it only because he could not have done otherwise;

(V) A person is morally responsible for failing to perform an action only if he could have performed it;

(VI) A person is morally responsible for an event only if he could have prevented it.

Blumenfeld (*op. cit.*) and Robert Cummins ("Could Have Done Otherwise," *The Personalist*, 60 (1979), 411-414) endorse III. Frankfurt (*op. cit.*) rejects III but endorses IV (thereby manifesting some tension in his opinion concerning the effect that showing P1—or P1'—false has on the traditional debate, insofar as he appears to think that he has thereby dissolved the debate but also resolved it by replacing P1' with the compatibilist IV—for I take it that such dissolution and such resolution are incompatible). Peter van Inwagen ("Ability and Responsibility," *The Philosophical Review*, 87 (1978), 201-224) endorses V and VI. And doubtless there are other principles which might be advanced in this vein as substitutes for P1'.

15. While *commentators* on Frankfurt's paper have ignored this claim, there is evidence that Frankfurt himself is at times strongly attracted to it. In the paper cited in note 3, for instance, he says (p. 830): "[T]here may be circumstances that constitute sufficient conditions for a certain action to be performed by someone and that therefore make it impossible for the person to do otherwise, but that do not actually impel the person to act or in any way produce his action." And in a later paper ("The Problem of Action," *American Philosophical Quarterly*, 15 (1978), 157-162), Frankfurt reiterates his claim that "the" principle of alternate possibilities is false, but now the principle to which he gives this name is P3 and not P1. However, there is a tension in the original paper (pointed out in note 14) concerning just what thesis it is for which Frankfurt wishes to argue. For, as will shortly be argued, if it is just P3 (or P3' and P7) that is (are) false, there is no need to think the debate between compatibilists and incompatibilists to be radically undermined (and, especially, no need to resort to principle IV of note 14, as Frankfurt does).
16. Cf. David Hume, *An Inquiry Concerning Human Understanding* (New York; the Bobbs-Merrill Co., Inc., 1955), p. 104; P. H. Nowell-Smith, *Ethics* (Harmondsworth, Middlesex: Penguin Books, Ltd., 1954), Ch. 19.
17. Cf. Gerald Dworkin, "Acting Freely," *Nous*, 4 (1970), 367-383; Harry G. Frankfurt, "Freedom of the Will and the Concept of a Person," *The Journal of Philosophy*, 68 (1971), 5-20; Gary Watson, "Free Agency," *The Journal of Philosophy*, 72 (1975), 205-220.
18. Cf. Thomas Reid, *Essays on the Active Powers of the Human Mind* (Cambridge, Mass.: the M.I.T. Press, 1969), pp. 254, 265; Richard Taylor, *Action and Purpose* (Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1966), Chs. 8-10.
19. More pedantically: if determinism is true, then there is no time  $t_1$  such that, for some person  $p$  and some action  $a$ ,  $p$  does  $a$  freely at  $t_1$ .
20. I take it that the incompatibilist will accept P4' because I take it that he will accept each of P8, principle I (see note 8), and the following principle:  
(VII) For any person  $p$ , any time and any action  $a$ :  
if  $p$  does  $a$  at  $t_1$  and can at  $t_1$  do other than  $a$  at  $t_1$  then  $p$  does  $a$  freely at  $t_1$ .  
Both I and VII (which is the converse of P3') are, I think, true and, when conjoined with P8, yield P4'.
21. Yet another position is captured in the following claim: (vi) the modified case of Peter, Paul, and the locked door shows P1' false, but it thereby shows only P7 false and it does not thereby show either P2' or P3' false. While I think it plausible to maintain that P7 is more obviously false than P3', I find it hard to come up with a plausible rationale for claim (vi), for I find it hard to come up with a plausible reason to maintain that P1' is false but both P2' and P3' are true.
22. An earlier version of this paper was read at the May 1981 meeting of the New Jersey Regional Philosophical Association. I am grateful for comments given on that occasion by Raziel Abelson, Lowell Kleiman, and Peter Klein.