



Alternative Possibilities in Frankfurt-style Cases

Marco Giovanni Stucchi

Abstract. In this article I will criticize the so called Frankfurt-style cases. These cases have been built with the purpose of denying that in order to hold a person responsible, it's necessary this person could have done otherwise, namely there were alternative possibilities. I present here two arguments according to which Frankfurt's cases did not succeed in this attempt. In the first argument I will focus, from and ethical and a logical point of view, on the Principle of Alternative Possibilities, and I will show that the Frankfurt-style cases are not truly able to falsify it. The second argument is semantical and concerns the formulation of Frankfurt example.

Keywords. Free Will, Harry Frankfurt, Liberty, Responsibility.

1 Introduction to the problem

It seems obvious that in order to hold a person responsible for something, we need to assume that she could have done otherwise. In other words we naturally take for granted the so called Principle of Alternative Possibilities (PAP): a person is morally responsible for what she has done only if she could have done otherwise (see Frankfurt (1969, p. 829)). This principle was accepted not only by the common sense, but also from the majority of philosophers, at least until Frankfurt challenged it in his very influential “Alternative possibilities and Moral Responsibility”. That paper, according to many, has proved that PAP is false. In this paper I would like to criticize Frankfurt’s argument by showing, first of all, that PAP must be clarified and, thereafter, that there are alternative possibilities in Frankfurt-style cases.

I will build here a Frankfurt-style case that should undermine PAP. In the literature there have been many complications and reformulations of Frankfurt’s original case in order to defend Frankfurt-style cases from some criticisms (see Mele and Robb (1998, p. 101) and see Fischer (2011, p. 249)). I believe however that the following example can be representative of the debate in general, and can work for the purpose of this paper. Suppose Mr. Black knows that Jones is a political dissenter and he desires an armed revolution. Black knows that tomorrow morning Jones will show up in front of the parliament with a gun in his coat in order to kill the Prime Minister¹. Black, however, is aware that Jones is kind of a coward man, and it would not be the first time that he show up armed in front of the parliament without doing anything. This time Black, who wants the death of the Prime Minister, has placed in Jones’ brain, without Jones knowing, a sophisticated mechanism, able to screen off Jones’ thought, and able to force him to kill the Prime Minister if Black pushes a certain button. Black does not want to interfere unnecessarily: he will push the button only if he sees that Jones shows the intention (and Black can check this “mental process” thanks to the mechanism) to not kill the Prime Minister. However that morning Jones is particularly resolute to kill the Prime Minister and, on his own, he shoots her with the gun. If, as Frankfurt persuasively suggests, Jones is morally responsible it seems that PAP is false, since he cannot do otherwise. In any case, in fact, he would have killed the Prime Minister.

Most philosophers have found Frankfurt-style cases compelling (Ginet 1996, p. 404), and this kind of argument has been used by the supporters of compatibilism about free will. Although many philosophers have taken for granted that Frankfurt has shown PAP to be false, there are some dissenters (Widerker 1995, p. 247 and see O’Connor 2005). There are different ways to attack Frankfurt-style

¹In the example at stake the agent is blameworthy, because I assume that the action is morally wrong. In this paper I will not consider whether Frankfurt-style cases can work both for blameworthy and praiseworthy actions (Nelkin 2011, p. 1), and I will assume a symmetry between these two.

cases; in this paper I will refer to some elements of Widerker's criticism, one of the most influential voices conflicting with Frankfurt.

2 Alternative possibilities in Frankfurt-style cases

I believe that in order to criticize Frankfurt-style cases we should look carefully at PAP, trying to understand what such a principle is supposed to mean. PAP as stated above, narrows down the responsibilities just for actions. But since, as van Inwagen points out (Fischer 2011, p. 245), we can hold individuals morally responsible also for others items such as omissions, consequences and intentions, I would like to generalize PAP. In this new formulation PAP claims that a person is morally responsible for S, only if the person could have avoided S. I will take PAP as divided in two different parts. The first part (*i*) is: "a person is morally responsible for S" and the second part (*ii*) is: "the person could have avoided S". These two parts are connected by a classic material conditional: $(i) \rightarrow (ii)$. If a person could have avoided S it means that is possible that S does not occur. Moreover, if it possible that S does not occur the person could have avoided S². If we formalize this we have $(ii) \leftrightarrow \Diamond \neg S$. Putting these elements together we have that if $(i) \rightarrow (ii)$ and $(ii) \leftrightarrow \Diamond \neg S$, it will follow that $(i) \rightarrow \Diamond \neg S$.

Consider now the example of the Frankfurt-style case in the first section. In that case we can substitute S with killing the Prime Minister. Frankfurt claims to have found a case in which $\neg \Diamond \neg S$, namely in which it is not possible that Jones does not kill the Prime Minister, but in which (*i*) is true, because we consider Jones responsible for killing the Prime Minister. In other words Frankfurt has shown that $(i) \rightarrow \Diamond \neg S$, which follows directly from PAP, is wrong.

I would like to take a moment to analyse what S means in this case. Just above we have taken S for 'killing the Prime Minister'. So at the question about what Jones is responsible for, it seems *prima facie* plain that the correct answer would be, simply, an action. If we say that S, the item for which Jones is hold responsible, corresponds with the action A 'killing the Prime Minister', it seems that Frankfurt is right. He actually gave us a case in which $(i) \rightarrow \Diamond \neg S$ is false. In fact, if we substitute A in this conditional we have that Jones is responsible for A, that is for his killing the Prime Minister, and that $\neg \Diamond \neg A$, since in any case, whether on his own or for the interference of Black, he would have killed the Prime Minister.

If it is just the action at stake then Frankfurt is right. But are we sure that Jones is responsible just for his action? It seems that in the example there is another item that plays a relevant role in Jones' responsibility. In my made-up case Jones does not just kill the Prime Minister, but he kills her with the intention to do it.

²I believe that I can claim this without claiming that the person must be responsible for her avoiding S.

Jones decides³ to perform a certain action, and he acts on the basis of this decision. With a closer look to my example we could say that S is the action A based on the intention Υ to perform such an action. Can Frankfurt's argument work if $S = A + \Upsilon$? In other words is it true that Jones is responsible for the action based on the intention to perform that action, despite the fact that it was impossible that his action based on the intention to perform that action does not occur? To find out let's substitute S with $A + \Upsilon$. I am assuming, and it seems reasonable to do it, that Jones is responsible for $A + \Upsilon$, so the antecedent of $(i) \rightarrow \Diamond \neg S$ is true. As we already know, in order to succeed in her aim a Frankfurt's supporter should prove that the consequent of our conditional is false, namely that $\neg \Diamond \neg S$. Since I am examining the case in which S is $A + \Upsilon$, it is clear that Frankfurt's supporter needs to prove that $\neg \Diamond \neg (A + \Upsilon)$. Frankfurt's supporter has an easy time in doing so because she can argue in the same way we have seen above. Since there is no possibility that the action does not occur, because of the presence of Black's mechanism, there is no possibility that the action based on the intention to perform that action would not occur. Frankfurt's supporters can claim in a convincing way that $\neg \Diamond \neg (A + \Upsilon)$, and therefore that PAP is wrong.

We have seen two cases so far: one in which S is identical to A and one in which S is identical to $A + \Upsilon$. In the first Jones is morally responsible for killing the Prime Minister, in the second Jones is morally responsible for killing the Prime Minister based on his intention for acting in this way. In these two cases we have seen that Frankfurt's argument is successful. I suggest however that now we should test carefully our intuitions about what Jones is responsible for. It seems sensible to affirm that Jones is responsible for his acting in such a way with the intention to do it (our second case). I suggest now to divide the two items at stake in this case: the action A and the intention Υ . It could be, in fact, that it is just one of these two items that plays a relevant role in holding Jones responsible. I suggest the following device in order to identify the item we are looking for: the item essential for Jones' responsibility (Mele 2008, p. 277). In this situation we have three elements: (i) Jones' responsibility, A Jones' killing the Prime Minister, Υ Jones having the intention to kill the Prime Minister. I have already established that if both A and Υ occur we have (i) as an outcome. Now I am going to see if (i) depends just on one of these two items. If I verify that removing one of these two items, and leaving the other as it is, there is a change in the truth value of (i), I can claim that (i), Jones' responsibility, depends on the former item; therefore I will conclude that such an item plays a fundamental role in Jones' responsibility.

³Although it could be useful, at a deeper level, to distinguish between 'have the intention to perform an action' and 'decide to perform an action', I believe that, in regard to this paper, this distinction is not necessary.

⁴I will use the non-logical sign + without any precise meaning. $A + I$ means just the action based on the intention to perform such an action. If I have used the logical connective \wedge , I would have run into some complications, as the reader can certify.

Take the case in which Jones, that morning, was overwhelmed by his same old cowardice, and he abandoned his intention to kill the Prime Minister. What would happen in this case? Black, who can read Jones' mind as a book thanks to his mechanism, would interfere and make Jones kill the Prime Minister. What should we say about Jones' responsibility in a such a situation? It seems obvious that, if we know the circumstances, namely that in Jones' brain there is a mechanism that will force him in the murder if he does not show the intention to kill the Prime Minister, nobody would hold Jones morally responsible for his killing the Prime Minister. If we remove Jones' intention to kill the Prime Minister (i) is false. On the other hand, according to our example, given that Jones has the intention to kill the Prime Minister, we would hold, as Frankfurt clearly states, Jones responsible for his killing the Prime Minister. The point seems to be clear: if there is Jones' intention to kill the Prime Minister then Jones is responsible ($\Upsilon \rightarrow (i)$); if there is no intention to kill the Prime Minister then Jones is not responsible ($\neg\Upsilon \rightarrow (i)$). From this follows that $\Upsilon \leftrightarrow (i)$. What should we say, in the Frankfurt-style case at issue,⁵ about the role of the action A in the explanation of Jones' responsibility? I have shown two cases in which (i)'s truth value changes: the case in which there is Jones' intention to kill the Prime Minister, and we have that (i) is true, and the case in which Jones has no intention to kill the Prime Minister, and we have that (i) is false. In both cases A occurs. It seems evident therefore, that the action does not play an essential role in Jones' responsibility, while the intention does, since, as I have shown, $\Upsilon \leftrightarrow (i)$.

I believe I have just proved that S, the item for which Jones' is morally responsible, corresponds with Υ , namely Jones' intention to kill the Prime Minister. Therefore PAP, in my case, takes this shape: $(i) \rightarrow \diamond\neg\Upsilon$. To reiterate, in order to prove the validity of Frankfurt's criticism to PAP, it must be proved that the consequent is false. Hence, in this last case, Frankfurt's supporter should prove that $\neg\diamond\neg\Upsilon$. But this, as we have just seen, is not the case. About Υ there are alternative possibilities: both $\diamond\neg\Upsilon$ and $\diamond\Upsilon$ are true. Jones has the possibility to have the intention to kill the Prime Minister and the possibility not to have such an intention. The interference system prevents Jones only from the possibility of $\neg A$, because in any case Jones will act A, namely $\neg\diamond\neg A$. Frankfurt's supporter is not able to show that $(i) \rightarrow \diamond\neg\Upsilon$ is false. I have proved that, as regarding the item that truly plays a role in my case, there are alternative possibilities. These alternative possibilities are relevant for responsibility because, if my argument works, they concern the essential item that explains responsibility itself.

⁵Which I take as model for the most common Frankfurt-style cases.

3 A criticism to my argument and another proposal

The main argument of this paper has been already presented in the second section. In this section I would like to show a possible criticism to my argument and to address to it, and finally, I will sketch a new argument against Frankfurt-style cases that should prove what I have already shown in the second section.

A Frankfurt supporter can raise an objection to my argument. She can build a different example in which Black's interference can happen in a different moment. In my previous example Black pushes the button after Jones shows an intention about his action. If he shows an intention to kill the Prime Minister Black will not interfere; on the other hand if he shows an intention to not kill the Prime Minister Black will interfere. The moment in which Black can interfere, removing alternative possibilities, is after the intention. Frankfurt's supporter could make up a case in which Black can interfere before the intention. This move should allow Frankfurt's supporter to remove the alternative possibility that I discovered in section 2. Let's make explicit the example I am talking about. Jones is the same person I described above, and he shows up in front of the parliament with a gun. The night before Black had installed a special mechanism in Jones' brain. This mechanism is able to show if Jones is *forming* the intention to kill the Prime Minister or not, and, in the case that Jones is forming the intention not to kill the Prime Minister, Black can push a button that triggers a process in Jones' brain, that ends with Jones' intention to kill the Prime Minister. That morning in front of the parliament Jones forms on his own the intention to kill the Prime Minister and he acts according to his intention. In this case it seems that we hold Jones responsible for his killing the Prime Minister, and there was not an alternative possibility about his intention to do it. Black's mechanism prevents Jones from the intention not to kill the Prime Minister. If we come back to the logical formalization I proposed in section 2, it means that, given that S is Υ , (i) is true and $\neg\Diamond\neg\Upsilon$. In this case $(i) \rightarrow \Diamond\neg\Upsilon$ is false, therefore Frankfurt is right about PAP.

I think that this criticism to my argument can be rejected (see Widerker (1995, pp. 249-251)). I suggest thinking about what 'forming an intention' really means, since forming an intention plays an important role in the last example. I believe that 'Jones is forming an intention' can have two different meanings relevant for my inquiry. Jones is forming an intention Υ can mean either 1) Jones will have the intention Υ ⁶ or 2) Jones will likely have the intention Υ but he can change his mind. It seems to me that these two cases are jointly exhaustive. Take the case 1). It seems that in this case there is no difference with the case I have analysed in section 2. If forming the intention will necessarily bring about the in-

⁶It this case there is a necessary relation between forming an intention and having such an intention.

tention, Frankfurt's supporter has just pushed the debate one step back without any change regarding the outcome. In fact I can just move the responsibility from the intention to the forming the intention,⁷ and I can show that in this last example there are indeed alternative possibilities about forming the intention. Let's move on to 2). In this case it seems obvious, by definition, that Jones has alternative possibilities: he can form his intention to kill the Prime Minister, but still he could change his mind and give up that intention. If what I have just said is coherent I see no reason why my argument against Frankfurt-style cases should be undermined by this last more sophisticated example.

I would like now to mention another possible way to attack Frankfurt-style cases. I will not go into detail with this part, because I believe I have already proved what was of concern, namely the fact that in Frankfurt-style cases there are alternative possibilities about the item relevant for the moral responsibility.

PAP states that an agent is morally responsible for what she has done only if the agent could have done otherwise. I believe that Frankfurt-style cases do not prove that PAP is wrong since they change the meaning of the term 'agent' in PAP. More precisely I think that when Frankfurt's supporter claims that "an agent is morally responsible for what she has done, even if she could not have done otherwise", 'agent' and 'she' do not have the same meaning since 'she' is not an agent at all. Frankfurt's supporter needs to prove that the individual that is morally responsible for something is the same individual that could not have done otherwise. I believe that as regarding (*i*), namely holding responsible Jones, we are holding responsible an agent, in the genuine sense of agent. But what should we say about (*ii*)? By saying that Jones does not have alternative possibilities, are we talking about an agent in the genuine sense of agent? If the answer is negative, as I would like to suggest, it seems that Frankfurt's supporter, in her attack to PAP, is playing with words since the subject of (*i*) is not the same subject as in (*ii*). Saying that Jones has not alternative possibilities, in other words that he could not do otherwise, means, in the case at stake, that Jones cannot not kill the Prime Minister. Jones cannot not kill the Prime Minister because Black's mechanism, by brainwashing Jones' mind, would force him to kill the Prime Minister if he shows the intention not to kill the Prime Minister. If we say that Jones has no alternative possibilities, we are saying that in the possible world in which Jones does not show the intention to kill the Prime Minister, Black would push the button, causing Jones' to kill the Prime Minister. I believe that in this possible world Jones, since he would be brainwashed, would lack his property to be an agent because he would lack any kind of control that we expect from an agent.

⁷If we assume that not only forming an intention Υ necessarily implies having an intention Υ , but also that, as it should be plain, having an intention Υ necessarily implies that one had forming an intention Υ ; there is, as a matter of fact, no difference between forming an intention Υ and having an intention Υ .

If the reader shows some resistances regarding what I have just argued, I think that this it is due to the particular story of the example. In my example Jones, that morning, goes out with a gun in his coat and with hate for the Prime Minister in his heart. If he will not shoot, it is just because of his cowardice. Moreover we know that his deciding to kill the Prime Minister is up to him, and it is a matter of subtle feelings and thoughts in his mind. Maybe we could be inclined to see the interference mechanism as a little push in a certain direction. I would like to change the story in order to give force to my argument.

That morning, in front of the parliament, there will be a parade in which the Prime Minister will walk among the people and she will shake some hands. Also in this example Black hates the Prime Minister, but not so much as in the previous one. Black decides to put (on the sly) a mechanism in the brain of a random citizen, who will participate at the parade. This mechanism shows to Black if the citizen will have the intention the punch the Prime Minister in the face, and it will enable Black, if the citizen does not show such an intention, to press a button that will cause the citizen to punch the Prime Minister. We do not know anything about this citizen, but as it normally happens, we would not think that she is going to punch the Prime Minister. As expected, the unlucky citizen does not show the intention to punch the Prime Minister, therefore Black interferes. Would we say that the citizen, in this case, preserves her property to be an agent? I do not think so. She is totally forced to do something nobody would ever expect from her. If I am right Frankfurt's supporter is not able to prove that the agent, properly speaking, has no alternative possibilities. The subject of *(ii)* for whom there are no alternative possibilities, is not an agent as the subject of *(i)* that we hold responsible. If I am right Frankfurt's supporter has not proved that PAP is false.

4 Conclusion

I think more should be said about this last criticism against Frankfurt-style cases. First of all the notion of agent could be discussed in details. However I believe that I have already carried out the task of this paper. In conclusion I would like to note that even if Frankfurt-style cases do not represent a valid argument for compatibilism, there can be still some good reasons for compatibilism. For example one could conceive liberty in the Kantian negative sense, namely as the freedom from specific bounds and interferences. The reader can easily note that, according to such conception, liberty can survive the loss of PAP.

References

- Fischer, John Martin (2011). "Frankfurt-Type Examples and Semicompatibilism: New Work". In: *The Oxford Handbook of Free Will (second edition)*. Ed. by Robert Kane. Oxford: Oxford University Press, pp. 243–265.
- Frankfurt, Harry (1969). "Alternate Possibilities and Moral Responsibilities". In: *Journal of Philosophy* 66.1, pp. 829–839.
- Ginet, Carl (1996). "In Defense of the Principle of Alternative Possibilities: Why I Don't Fine Frankfurt's Argument Convincing". In: *Philosophical Perspectives* 10, pp. 403–417.
- Mele, Alfred (2008). "Manipulation, Compatibilism and Moral Responsibility". In: *The Journal of Ethics* 12.3, pp. 263–286.
- Mele, Alfred and David Robb (1998). "Rescuing Frankfurt-Style Cases". In: *The Philosophical Review* 107.1, pp. 97–112.
- Nelkin, Dana Kay (2011). "Abilities". In: *Making Sense of Freedom and Responsibility*. Ed. by Dana Kay Nelkin. Oxford: Oxford University Press.
- O'Connor, Timothy (2005). "Freedom with a Human Face". In: *Midwest Studies in Philosophy* XXIX, pp. 207–227.
- Widerker, David (1995). "Libertarianism and Frankfurt's Attack on the Principle of Alternative Possibilities". In: *The Philosophical Review* 104.2, pp. 247–261.