



Philosophy Project

A.S.W.
Laksumanage



**SUPERVISED BY HAZEL (OXFORD
PHILOSOPHY & THEOLOGY GRADUATE AND
MASTER OF PHILOSOPHY IN
ANTHROPOLOGY)**

Taking Frankfurt by the Horns – What can Frankfurt Cases tell us about our moral responsibilities under determinism?

A.S.W. Laksumanage

Abstract

In our contemporary understanding of justice, we presume that individuals can only be held responsible for that which they have control over. This is the basis of 'intent' in our legal systems, and underpins a large portion of the ways in which we assign praise and blame on a daily basis. With the growing popularity of determinism, compatibilists like Harry G. Frankfurt sought to create a moral defence of praise and blame even under a deterministic worldview. This led to the creation of what are now known as 'Frankfurt Cases', instances in which an individual can be held responsible even if they lacked alternatives.

In this paper, I will attempt to go over current criticisms of Frankfurt-Style Cases (FSCs) and evaluate the possible shortcomings within those criticisms by analysing the responses made by defenders of Frankfurt. I will then attempt to propose my own criticism of FSCs and respond to possible criticisms of my own criticism levied by the pro-Frankfurt camp.

Before that however, it is necessary that we understand the importance of this discussion and what the implications of our conclusions are. If we come to the conclusion that PAP is necessary for one to be considered morally responsible, this has major implications for a worldview that is deterministic; ultimately, this would result in a world where no person is responsible for the actions they commit. This would render our current understanding of praise and blame almost null, and leave very serious doubt about our current understanding of morality and justice. In turn it would render our current systems of enforcing justice, such as the legal system, to not only be ineffective in providing justice but also responsible for historically condemning individuals who did not deserve punishment to face said punishment. The consequences of this conclusion therefore have extremely large implications for our world at large, and should not be taken lightly.

Literary Review

It seems intuitive that people should only be responsible for actions over which they have control: If I am forced to do an action, or am otherwise coerced into doing so, it seems reasonable that I should not be held morally responsible for that action on the basis that I could not do otherwise. This is known as the Principle of Alternate Possibilities (or PAP) and can be formalised as follows:

PAP: a person is morally responsible for what she has done only if she could have done otherwise.

This proposition was widely accepted up until the work of Harry G. Frankfurt who, in his 1969 paper 'Alternate Possibilities and Moral Responsibility', produced examples which we now refer to as 'Frankfurt-Style Cases'. These are cases in which Frankfurt claims an agent could still be considered morally responsible for their actions even if they lacked the ability to do otherwise (Frankfurt, 1969).

There have also been many critics of Frankfurt-Style Cases, mainly falling into one of two positions;

Position 1 – PAP as stated above is false but an adaptation of PAP or some neighbouring principle is immune to Frankfurt-Style Cases and therefore serves a similar purpose in showing that moral responsibility requires a degree of possible alternative.

Position 2 – Frankfurt-Style Cases fail to adequately disprove PAP and therefore moral responsibility is still dependent on an agent's ability to have done otherwise.

I will first examine Frankfurt's original works to establish a clear understanding of what it is I aim to oppose through this paper. In his 1969 paper, Frankfurt refers to

'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. A person may do something in circumstances that leave him no alternative to doing it, without these circumstances actually moving him or leading him to do it-without them playing any role, indeed, in bringing it about that he does what he does.'

What Frankfurt refers to here are instances where an agent truly has no ability to have done otherwise (i.e. instances where the actions they commit would have been committed under any circumstance) but yet still bears some degree of moral responsibility for the action. Frankfurt therefore argues that it is still possible for us to assign moral praise or blame to an actor regardless of their ability to have done otherwise. The most famous example of a Frankfurt case, surprising as it may be, is not one that Frankfurt himself wrote. Rather it comes from a 1982 paper by J. M. Fischer entitled "Responsibility and Control" in which he writes the following;

"Black is a nefarious neurosurgeon. In performing an operation on Jones to remove a brain tumor, Black inserts a mechanism into Jones's brain which enables Black to monitor and control Jones's activities. Jones, meanwhile, knows nothing of this. Black exercises this control through a computer which he has programmed so that, among other things, it monitors Jones's voting behavior. If Jones shows an inclination to decide to vote for Carter, then the computer, through the mechanism in Jones's brain, intervenes to assure that he actually decides to vote for Reagan and does so vote. But if Jones decides on his own to vote for Reagan, the computer does nothing but continue to monitor-without affecting-the goings-on in Jones's head.

Suppose Jones decides to vote for Reagan on his own, just as he would have if Black had not inserted the mechanism into his head. Then Frankfurt claims that Jones is responsible for voting for Reagan, regardless of the fact that he could not have done otherwise." (Fischer, 1982)

With Frankfurt's original argument thoroughly established, we must consider the most prominent objections to FSCs.

Objection 1 – FSCs do not eliminate the ability to have done otherwise

The first group of objections claim that FSCs are not truly instances in which an actor lacks the capacity to do otherwise.

Objection 1 falls into the first position of objections mentioned above, and claims that we can alter our understanding of PAP in a way that makes clear what the original intention of PAP was. Proponents claim that there is a distinction between 'general' abilities and the abilities one possesses under a particular counterfactual. Therefore, proponents claim that despite an actor's technical inability to have not done otherwise in a FSC, this does not extend to other instances where an individual is not externally coerced (Campbell, 1997).

Proponents thus choose to replace PAP with a new principle known as PAP-general, as below:

PAP-general: a person is morally responsible for what she has done only if she had the general ability to do otherwise.

The conclusion of this objection is that, in the instance of a Frankfurt case where an actor performs an action of their own accord, they still possess the ability to have done otherwise. It is merely true that, in the counterfactual where they choose to act otherwise, there is an external force that prevents them from doing so. In this way, the understanding of the 'inability to have done otherwise' in a FSC falls short as it conflates two different scenarios as the same.

Evaluation of Objection 1

This objection seems to be quite strong as it shows that FSCs do not actually prove that which they seek to prove while also showing that FSCs are not valid when generalised to other instances in which we make decisions. In doing so, it allows us to continue using PAP as our benchmark for moral praise and blame.

However, it may face a potential problem in terms of the distinction between 'general' abilities and abilities under a particular situation. This distinction still seems to be quite arbitrary as it doesn't necessarily prove why we should care about general abilities more than abilities in a specific circumstance. A defender of FSCs may very well respond that it doesn't matter what our general abilities are, and that all that matters is that there are 'potential' instances in which we can assign praise or blame without the ability to have done otherwise and that that is enough to refute PAP.

Nevertheless, it still seems reasonable to draw a distinction between the scenario in which the actor votes for Carter because they want to and the scenario in which they wanted to vote for Carter but were instead forced to vote for Reagan, and the differing moral responsibility between the two cases still applies.

Objection 2 – FSCs aren't 'realistic' enough

This objection falls into the first category of positions in opposition to FSCs. Proponents claim that the examples used in FSCs are too detached from reality and require a large degree of 'artificial' influence in order to occur (Stump, 1990). They supplement this with a new principle referred to as 'PAP-Ordinary'

PAP-ordinary: in ordinary circumstances, a person is morally responsible for what she has done only if she could have done otherwise.

Proponents of this objection thus claim that, in most circumstances of daily life we are able to apply PAP as an adequate means of determining praise and blame.

Evaluation of Objection 2

The shortcomings of this objection seem to be quite glaring, in that the idea of an 'ordinary circumstance' seems extremely subjective and vague. What allows us to consider a situation ordinary versus extraordinary? There is no existing criteria, or at least not one that is widely accepted, to distinguish a meaningful difference between these situations.¹

This objection also seems to be begging the question. It first assumes that Frankfurt Cases are 'not ordinary' and then asserts that PAP applies only in ordinary cases. In other words, 'PAP does not apply in instances where PAP does not apply'. It is for this reason that this objection fails to adequately respond to FSCs and is possibly the weakest of the objections I will consider in this paper.

Overall, this objection seems quite weak and doesn't genuinely tackle the ideas underpinning a FSC, rather simply claiming that we should not be able to engage with hypotheticals that we deem to be 'unrealistic'. In fact, this undermines the whole point of hypotheticals generally. In dealing with dilemmas such as those of FSCs, we are compelled to contend with challenging circumstances in order to reveal potential conclusions about our choices generally. We cannot then simultaneously claim that a hypothetical shouldn't be considered on the basis that it is 'unrealistic' the moment where it becomes too challenging for the current scope of our reason. If we seek a consistent set of criteria by which we assign praise and blame, we must also ensure those criteria apply in the most fringe of cases, or they aren't consistent criteria at all.

Objection 3 – The 2-Horn Dilemma

The third objection to FSCs states that there is an inherent tension within FSCs that means that they cannot respond to or undermine PAP in the way that Frankfurt claims that they do. This objection thus falls squarely under the second position of criticisms to FSCs.

Proponents argue that FSCs cannot be constructed in a way that means an agent truly lacks alternatives while continuing to assign blame to the aforementioned agent at the same time. To illustrate this, let us re-examine the example of the nefarious neurosurgeon. In his 1996 paper, Carl Ginet uses the difference in time at which the two scenarios occur to show that there is a difference between the version of an action committed without the intervention of Black (committed at time t1) and the version of the action committed with Black's intervention (at time t3) (Ginet, 1996).

To be more clear, Ginet posits that we should not consider the 2 instances of a FSC as the same in terms of their moral weight. Rather, we can consider the 2 possible scenarios to be different from one another. For example, in the case of the nefarious neurosurgeon and the microchip, we can consider the 2 scenarios to be of different moral weight. In the instance where Jones chooses to vote for Reagan without interference from Black's machine, we can consider him morally responsible for that action as he has engaged in said action while the alternative of acting in a different manner (voting for Carter) is

¹ In fact, it may even be correct to say that the definition of 'ordinary circumstance' will always be a subjective one. Instances which may be considered ordinary to one person may be extraordinary to another. Furthermore, our understanding of 'ordinary' circumstances may change over time. So while it may not be the case that FSCs are commonplace today, they may become commonplace with the development of new technologies such as Artificial Intelligence. One area to consider may be those of self-driving cars and the moral decisions one may have to undertake.

still theoretically open to him. However, in the instance where Jones instead would have voted for Carter but was compelled to vote for Reagan by Black's machine, Jones does not bear a similar moral responsibility as his action was compelled by an outside force.

It is for this reason that proponents of FSCs face a major dilemma; they must simultaneously (1) construct a scenario in which an agent truly lacks any alternative while considering the 2 scenarios as occurring at different times and (2) construct a scenario in which an actor maintains the same degree of moral responsibility. There seems to be no current means by which proponents of FSCs can achieve (1) without (2) or vice versa, and therefore FSCs fall short.

Evaluation of Objection 3

Defenders of FSCs would argue that, in the actual scenario of a FSC, there is no intervention from the outside force and that the actor does indeed engage in the action. Defenders of FSCs would conclude that any factor causally isolated from the actor cannot be relevant to assigning moral praise and blame. Thus, even if Jones isn't responsible, this isn't entailed by his lacking alternatives.

In retort, I feel that this response fails to address the problems brought up by the 2-horn dilemma. It fails to consider that it is possible for the 2 instances of the action being carried out to be considered morally distinct from one another in the same way we would in other instances. For example, a proponent of the 2-horn dilemma objection would claim that there is a morally relevant difference between someone being forced to lie upon pain of death and someone choosing to lie of their own accord. Thus the 2-horn dilemma stands up quite well to this initial criticism.

Overall, the 2-horn dilemma seems quite strong due to its conclusion that there is a fundamental tension within the construction of FSCs that do not allow them to achieve both of the criteria they wish to do simultaneously, which can allow us to conclude that a lack of alternatives and moral responsibility cannot exist simultaneously. This ultimately reinforces PAP, causing it to stand strong.

Overall Evaluation of current responses to FSCs

In evaluating the current objections to Frankfurt cases, there seems to be a key issue that FSCs face. Namely, that they seem to conflate two independent (and thus morally different) scenarios to be equivalent, and in doing so fail to adequately prove that we can continue to bear moral responsibility even in the instance we are coerced into committing an action. Therefore, it seems that PAP remains strong in terms of its bearing on whether we assign praise or blame.

Thus we can safely claim that PAP still remains a criteria for moral praise or blame, and can at the very least be adjusted in a way that reflects the original intention of PAP to show that a lack of alternatives and moral culpability are indeed mutually exclusive.

Impulse and Action - the problem of praise and blame under PAP

The next part of this paper will then claim that, in light of PAP, it is not possible for us to assign praise or blame to an individual. I will begin by laying out my argument in its entirety before defending the individual premises and conclusions in-depth.

1. In every instance an individual chooses to commit an action, they commit it as a result of one of two contributing factors;
 - (a)they are forced into committing said action
 - (b)they 'want' to commit said action
2. Even when we commit an action that we may not 'want' to commit, we do so as a result of a superseding desire(Impulse) which compels us to commit said action
3. In instance (a), under PAP the agent lacks moral responsibility for their action as they could not have acted otherwise
4. In instance (b), we commit an action because we act upon a 'want', either direct(going to a lecture because I want to do so) or 'Impulse'(going to a lecture because I want to avoid failing my current semester)
5. Our wants are produced internally of our brains, either due to biological needs and wants or as a response to external stimuli
6. We therefore cannot control that which we want
7. We thus conduct all of our actions on the basis of impulses or wants that we have no control over
8. Under PAP, the agent also lacks moral responsibility in instance (b) as they do not have control over their action and thus could not have acted otherwise
9. There is therefore no instance under PAP in which an agent can be held morally responsible for their actions

I will now begin my defence of this argument, starting with the 1st premise.

1. In every instance an individual chooses to commit an action, they commit it as a result of one of two contributing factors;
 - (a)they are forced into committing said action
 - (b)they 'want' to commit said action

Critics may claim but there are instances in which one commits an action where they are not forced yet do not 'want' to do so. For example, I may not want to go to my lecture but I do so anyway. In order to avoid confusion I will then make clear that when I use the term 'want' I am not using it in the colloquial sense- of enthusiastic desire. Rather I mean want purely in terms of an impulse that causes us to do an action. Thus I may not 'want' to go to my lecture but I still go to my lecture because I have an 'impulse' to do so.

This 'impulse' can also be considered a form of 'want' in itself, and in fact it is often a result of a superseding want or overriding desire that we may already possess. Perhaps I do not want to attend my lecture, but I have a superseding want to not fail my semester. The want to not fail is an overriding disposition that we act upon, yet it is still not one whose origins we have meaningful control over.

2. Even when we commit an action that we may not 'want' to commit, we do so as a result of a superseding desire(Impulse) which compels us to commit said action

Thus we have established that there are only 2 instances in which we choose to do something. I will now attempt to prove that we lack the ability to have done otherwise in both of these instances, and thereby show that we are unable to assign praise and blame.

3. In instance (a), under PAP the agent lacks moral responsibility for their action as they could not have acted otherwise
4. In instance (b), we commit an action because we act upon a 'want', either direct (going to a lecture because I want to do so) or 'impulse' (going to a lecture because I want to avoid failing my current semester)

Most would agree that in the instance one is forced to do an action they already lack moral responsibility for the action. I propose that this same lack of moral responsibility also exists in instances where we choose based on our wants.

5. Our wants are produced internally of our brains, either due to biological needs and wants or as a response to external stimuli
6. We therefore cannot control that which we want
7. We thus conduct all of our actions on the basis of impulses or wants that we have no control over
8. Under PAP, the agent also lacks moral responsibility in instance (b) as they do not have control over their action and thus could not have acted otherwise

The reason I claim this is that we have no meaningful way of controlling that which we want. Our wants and desires, despite being internal or produced within our brains, are still not within our control. Thus, we continue to conduct all our actions on the basis of impulses that we have no control over and in doing so we lack control over our own actions. Thus, we lack any alternatives to committing our actions and thereby cannot assign praise or blame even when we commit an action because we 'want' to. This applies in all instances of choice, be it as mundane as our choice of clothing or as important as our life choices regarding education or moving to a different country. In both instances we act upon our 'wants' and nothing more, and in both instances we are in some way acting upon impulses resulting from social conditioning.

9. There is therefore no instance under PAP in which an agent can be held morally responsible for their actions

Objectors to my claim may state that even if we possess desires, we have the ability to not act upon those desires or choose to act otherwise. For example, I may 'want' to go to my lecture but I still may not go. The problem with this response is that, in choosing to not act upon our wants, we make that decision on the basis of another want. Namely, the desire to not act upon our desires. This is simply another want that we do not have control over and therefore is again, outside our control. (Van Inwagen, 1980). We are trapped within an inescapable cycle of want that controls our every action.

Therefore, it is clear that as a result of PAP remaining strong (or at the very least a slight adjustment to PAP remaining strong) we are unable to assign praise and blame and therefore our current systems of assigning moral responsibility remain unsubstantiated. Perhaps there are other means by which we can assign praise and blame that do not require PAP, but proponents of this new principle would have to show why it doesn't also fall prey to the problem of action being determined by want. It would also have

to, at the same time, not lead to morally abhorrent scenarios where we can assign praise or blame to an individual who hasn't committed a morally praiseworthy or blameworthy action.

Possible Areas for Future Research

There may also be ways in which one could extend upon the conclusions raised in this paper. For example, whether there really is a meaningful difference between an impulse being 'internal' and 'within our control', and whether we should treat them as such. If so, this would have major implications for what we can hold individuals responsible for, such as our understanding of whether an individual with a mental illness that forces them to commit immoral actions is just as equally responsible for those actions as a person of sane mind. Studies on whether different cultures perceive the criteria for moral responsibility differently could be a possible avenue for understanding alternative means for assigning praise and blame. (Hannikainen, 2019)

Conclusion

In conclusion, it seems that the current objections to FSCs clearly prove that PAP stands strong. In doing so, they also show that we currently lack a basis for assigning moral praise and blame that is consistent with both our understanding of what entails moral responsibility as well as our understanding of what we are in control of. We must therefore radically transform the way in which we understand moral responsibility in order for us to continue the systems of praise and blame that are so central to our systems of morality, or we are left with the difficult decision of abandoning our current moral systems.

Bibliography

Campbell, Joseph Keim, (1997), "A Compatibilist Theory of Alternative Possibilities", *Philosophical Studies*, 88(3): 319–330.

Chisholm, Roderick M. (1967). "He Could have Done Otherwise". *The Journal of Philosophy*, 64(13), 409–417. doi:10.2307/2024211

Fischer, John Martin (1982). Responsibility and Control. *The Journal of Philosophy*, 79(1), 24–40

Fischer, John Martin (1994), *The Metaphysics of Free Will*, Oxford: Blackwell.

Fischer, John Martin (1999). Recent Work on Moral Responsibility. *Ethics*, 110(1), 93–139.

Frankfurt, Harry G. (1969). Alternate Possibilities and Moral Responsibility. *The Journal of Philosophy*, 66(23), 829–839

Ginet, Carl (1996). "In Defense of the Principle of Alternative Possibilities: Why I Don't Find Frankfurt's Argument Convincing". *Philosophical Perspectives*. 10: 403–417.

Hannikainen, Ivar R. (2019). "For Whom Does Determinism Undermine Moral Responsibility? Surveying the Conditions for Free Will Across Cultures". *Front Psychol.* 2019; 10: 2428.

Robb, David, "Moral Responsibility and the Principle of Alternative Possibilities", *The Stanford Encyclopedia of Philosophy* (Fall 2020 Edition), Edward N. Zalta (ed.)

Stump, Eleonore, (1990), "Intellect, Will, and the Principle of Alternative Possibilities", in M. D. Beaty (ed.), *Christian Theism and the Problems of Philosophy*, Notre Dame, IN: University of Notre Dame Press, pp. 254–85

Van Inwagen, Peter, (1980), "The Incompatibility of Responsibility and Determinism", in M. Bradie and M. Brand (eds.), *Bowling Green Studies in Applied Philosophy*, vol. 2, Bowling Green, OH: Bowling Green State University, pp.30–36

Widerker, David (1995). Libertarianism and Frankfurt's Attack on the Principle of Alternative Possibilities. *The Philosophical Review*, 104(2), 247–261