

A DILEMMA FOR FRANKFURT-STYLE CASES, AND THE DEFINITION OF
'ROBUSTNESS'

By

JACLYN GENCHI

A THESIS PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF ARTS

UNIVERSITY OF FLORIDA

2006

Copyright 2006

by

JACLYN GENCHI

TABLE OF CONTENTS

	<u>page</u>
ABSTRACT.....	iv
INTRODUCTION	1
AN OBJECTION TO FRANKFURT-STYLE CASES, AND REPLIES	12
The First Horn of the Dilemma: Assumption that the Prior Sign Determines the Agent's Choice (i.e., That the Prior Sign Deterministically Causes the Agent's Choice).....	20
The Second Horn of the Dilemma: Assumption that the Prior Sign is Not Sufficient for the Agent's Choice (i.e., That the Agent's Choice Is Indeterministically Caused)	29
No-Prior-Sign Cases.....	30
Blockage Cases.....	34
Internal-Sign Cases.....	40
Necessary-Condition Cases	43
THE DEFINITION OF 'ROBUSTNESS'	49
LIST OF REFERENCES	69
BIOGRAPHICAL SKETCH	71

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Master of Arts

A DILEMMA FOR FRANKFURT-STYLE CASES, AND THE DEFINITION OF
'ROBUSTNESS'

By

Jaclyn Genchi

August 2006

Chair: Jon Tresan

Major Department: Philosophy

The purpose of this thesis is to argue that one particular objection to Frankfurt-style cases has thus far not been shown to be mistaken. I defend the objection called the dilemma defense of the Principle of Alternate Possibilities objection from the replies given by proponents of Frankfurt-style cases to defend these cases from this objection.

INTRODUCTION

Frankfurt-style cases are traditionally taken to show that the Principle of Alternative Possibilities (PAP) is false¹ – to show that alternative possibilities are not required for moral responsibility. In this essay I will be discussing a particular objection to Frankfurt-style cases, the dilemma defense of the Principle of Alternate Possibilities. I will proceed by first explaining the role of Frankfurt-style cases in the debate concerning moral responsibility, and will follow with an explanation of the objection to Frankfurt-style cases I will be discussing, the dilemma defense of PAP. Proponents of Frankfurt-style cases have responded to both horns of the dilemma defense of PAP and, after reviewing these responses, I will argue that such responses are inadequate and, so, the dilemma defense of PAP stands as an objection to Frankfurt-style cases. I will give substantial attention to the strategy used by Derk Pereboom as a response to the dilemma defense of PAP, arguing that it fails.

Throughout this essay I will be discussing one of the many facets of the debate concerning moral responsibility, whether Frankfurt-style cases succeed in showing that the Principle of Alternate Possibilities is false. In considerations of moral responsibility, much of the debate has been focused on what is called the Compatibility Question: can we be morally responsible agents in a deterministic world? We are concerned to answer this question because there seems to be a conflict between our picture of ourselves as

¹ The Principle of Alternate Possibilities (PAP) is as follows: having an alternative possibility to choosing/doing x is necessary for an agent to be morally responsible for choosing/doing x .

agents that have control over what we choose and do, and our picture of ourselves as parts of a world governed by natural-law-abiding forces. We believe that what we do is in some way up to us, while at the same time we believe that what we do is a part of the causal order of the world, and these two pictures of ourselves seem to be in conflict. The thesis of determinism² is, roughly, that the events of the past together with the laws of nature entail the events of the present and future; if all events of the present and future are entailed by the past, it seems that all of our choices and actions are entailed by events (together with the laws of nature) that happened before we were born. If the thesis of determinism is true, we are challenged to find a way in which we can maintain our view that we have any control over our choices, our actions, the way in which we perform our actions, whether we perform our actions, etc.

Two positions are taken in response to the Compatibility Question: Compatibilism and Incompatibilism; the Compatibilist claims that moral responsibility and determinism are compatible (i.e., that the truth of the thesis of determinism would not entail that we cannot be morally responsible for what we do), while the Incompatibilist claims that moral responsibility and determinism are not compatible (i.e., that the truth of the thesis of determinism would entail that we cannot be morally responsible for what we do).

There are two approaches a Compatibilist position may take. On the one hand, a Compatibilist may claim that whether or not determinism obtains is irrelevant to whether or not we can be morally responsible for what we do. P.F. Strawson³ and Harry G.

² I like Peter van Inwagen's statement of the thesis of determinism in *An Essay on Free Will*. According to van Inwagen, the thesis of determinism is the conjunction of the following two theses: (1) For every instant of time, there is a proposition that expresses the state of the world at that instant, and (2) If p and q are any propositions that express the state of the world at some instants, then the conjunction of p with the laws of nature entails q [van Inwagen, 1983, p. 65].

³ Strawson, 1982.

Frankfurt⁴ both present Compatibilist views of this sort. On the other hand, a Compatibilist may claim that we can be morally responsible for what we do because whatever condition(s) is (are) necessary and sufficient for being morally responsible are compatible with determinism. Compatibilism of the latter sort has been the most prominent Compatibilist approach advocated in recent literature, notably by John Martin Fischer.⁵

There are three kinds of Incompatibilist: the Libertarian, the Hard Determinist, and the Hard Incompatibilist. The Libertarian claims that at least some of the things we do are not determined and, so, we can be morally responsible for such things, while both the Hard Determinist and the Hard Incompatibilist claim that we cannot be morally responsible for any of the things we do. The Hard Determinist claims (disagreeing with the Libertarian) that all of the things we do are in fact determined and, so, we cannot be morally responsible for anything that we do. The Libertarian and the Hard Determinist agree, though, on the following point: that a necessary condition on being morally responsible for some action is that it be not determined. This view of moral responsibility is generally taken to be *the* Incompatibilist intuition, the core of the Incompatibilist position. Traditionally, I believe, most Incompatibilists were led to either the Libertarian or Hard Determinist position by first believing that a requirement of my being morally responsible for something I do is that in so doing I was not determined to do so. There is, though, a variant of the Incompatibilist position, the Hard Incompatibilist, which adds that even an action that was not determined, but was undetermined, is not something that

⁴ Frankfurt, 1971.

⁵ Fischer, 1994.

we can be held morally responsible for. The Hard Incompatibilist claims that we cannot be morally responsible for anything we do, regardless of whether or not all of what we do is determined. The Hard Incompatibilist claims (disagreeing with both the Libertarian and Hard Determinist) that regardless of whether the thesis of determinism is true or false we are not morally responsible for what we do because we cannot be morally responsible for undetermined choices/actions as well as determined choices/actions. The Hard Incompatibilist claims that indeterminacy would not allow us the sort of control we believe is required for moral responsibility and which the Incompatibilist believes to be missing in a deterministic world.

There is another question (one related to the Compatibility Question) that is central to the debate concerning moral responsibility: what condition(s) must be satisfied in order for an agent to be morally responsible for something he has done? Intuitively, there is at least one condition that must be satisfied in order for an agent to be responsible for what he does: the agent must have some kind of control with respect to this choice/action. Both Compatibilists and Incompatibilists (for the most part) agree that an agent's having some kind of control over what he does is necessary for that agent to be morally responsible for so doing. There is major disagreement, though, over just what kind of control is necessary for being morally responsible. There are two main positions that are taken on this issue. One position claims that having the relevant sort of control amounts to an agent being the source or originator of his actions; one who takes this position is either a Source Compatibilist or a Source Incompatibilist. The Source Compatibilist holds that the sort of control associated with being the source or originator of one's actions is a sort of control that both is sufficient for moral responsibility and is compatible with

determinism, whereas the Source Incompatibilist holds that the sort of control associated with being the source or origination of one's actions is undermined by determinism.

Another position taken with respect to the question of what kind of control is required for moral responsibility is that this kind of control amounts to having and/or exercising the ability to choose a course of action from alternatives; one who takes this position is either a Leeway Compatibilist or a Leeway Incompatibilist. Both the Leeway Compatibilist and the Leeway Incompatibilist hold that what is (at least partly) relevant to an agent's being morally responsible for something is that he could have done otherwise than he actually did, that he had open to him alternative possibilities. Both the Leeway Compatibilist and the Leeway Incompatibilist adhere to the Principle of Alternate Possibilities (PAP): having alternate possibilities to what one does is necessary for being morally responsible for so doing. The Leeway Compatibilist holds that having alternatives to what one does is necessary for being morally responsible, and that having them is compatible with determinism. The Leeway Compatibilist will typically claim that to claim that one has alternatives to what one does is to claim that one would have done otherwise if one had so chosen, where this ability is taken as compatible with one's actual choice being determined. The Leeway Incompatibilist holds that the ability to choose/do otherwise is necessary for being morally responsible, and that this ability is undermined by determinism; if determinism is true, then one never really has any alternative possibilities to what one actually chooses/does. If determinism is true, then the past together with the laws of nature entails what one chooses/does at any moment and, so, if determinism is true, one never can do anything other than what one does do.

1. We can see that the truth of PAP is playing a crucial role in the Leeway Incompatibilist's argument for the Incompatibility of moral responsibility and determinism. The argument presented by the Leeway Incompatibilist for the incompatibility of moral responsibility and determinism is as follows:
2. Moral responsibility requires having alternate possibilities (PAP).
3. If the thesis of determinism is true, then one never can do otherwise than what one actually does, one never has any alternative possibilities.
4. Therefore, the truth of the thesis of determinism entails that no one is ever morally responsible for what one does (i.e., moral responsibility and determinism are incompatible).

Thus, we can see that if PAP is false, the Leeway Incompatibilist's argument for the incompatibility of moral responsibility and determinism fails.

In his 1969 essay "Alternate Possibilities and Moral Responsibility," Harry Frankfurt proposed that a case could be given to show that PAP is false. David Widerker gives a nice description of the sort of case that Frankfurt is supposing would show PAP false:

(FR) There may be circumstances in which a person performs some action which, although [these circumstances] make it impossible for him to avoid performing that action, [these circumstances] in no way bring it about that he performs it.⁶

If a case could be given in which (i) a person cannot do otherwise than what he does do (i.e., in which the agent has no alternate possibilities) and (ii) whatever explains why he cannot do otherwise is not what causes or brings about the action he performs, then if in such a case (iii) the agent is morally responsible for what he does, this would show that having alternate possibilities is not required for moral responsibility. If a case could be given in which the agent has no alternate possibilities to what he actually does (condition (i) of an FR scenario), and yet he is morally responsible for what he does (condition (iii) of an FR scenario), this would show that having alternate possibilities to what one does is

⁶ Widerker, 2002, p. 323.

not required for being morally responsible for so doing. In order to give a case that meets both conditions (i) and (iii) of an FR scenario, it must be that the thing that makes it the case that the agent cannot do otherwise is not the same thing that brings about the agent's choice or action (condition (ii) of an FR scenario), because if what removes the agent's alternate possibilities is the same thing that brings about the agent's choice or action, we would claim that agent was compelled or constrained to choose or act as he did, in which case he would not be morally responsible for so doing. If a case could be given that meets all three conditions of an FR scenario, then such a case would be a counterexample to PAP. The following case, which is John Martin Fischer's standard example, is a Frankfurt-style case, a case meant to satisfy all three conditions of an FR scenario:

Suppose Jones is in a voting booth deliberating about whether to vote for Gore or Bush. After reflection, he chooses to vote for Gore and does vote for Gore by marking his ballot in the normal way. Unbeknownst to him, Black, a liberal neurosurgeon working with the Democratic Party, has implanted a device in Jones's brain which monitors Jones's brain activities. If he is about to choose to vote Democratic, the device simply continues monitoring and does not intervene in the process in any way. If, however, Jones is about to choose to vote, say, Republican, the device triggers an intervention that involves electronic stimulation of the brain sufficient to produce a choice to vote for the Democrat (and a subsequent Democratic vote). How can the device tell whether Jones is about to choose to vote Republican or Democratic? This is where the "prior sign" comes in. If Jones is about to choose at T_2 to vote for Gore at T_3 , he shows some involuntary sign – say a neurological pattern in his brain – at T_1 . Detecting this, Black's device does not intervene. But if Jones is about to choose at T_2 to vote for Bush at T_3 , he shows an involuntary sign – a different neurological pattern in his brain – at T_1 . This brain pattern would trigger Black's device to intervene and cause Jones to choose at T_2 to vote for Gore, and to vote for Gore at T_3 .⁷

So, in this case, Jones does show the relevant prior sign and does vote for Gore, without any intervention from the device. It seems, then, plausible to claim that Jones is morally responsible for his choice to vote for Gore. But, had Jones not shown the relevant

⁷ Fischer, 2002, p. 282.

prior sign (i.e., had Jones shown a sign indicating that he would do otherwise than choose to vote for Gore), the device would have intervened and caused Jones to choose to vote Gore. So, this shows that Jones could do not do otherwise than choose to vote for Gore; in this case, Jones has no alternate possibility to his actual choice. This case, then, seems to be an FR scenario, since this is a case in which (i) Jones has no alternate possibility to what he does (where “does” in this case refers to both his choice to vote for Gore and his subsequent action of voting), (ii) whatever makes it the case that Jones has no alternate possibility - in this case the intervening device - in no way brings about Jones’s actual choice, and (iii) Jones is morally responsible for what he actually does. Since all three conditions of an FR scenario seem to be satisfied by Fischer’s case, this Frankfurt-style case seems to be a counterexample to PAP.

Before any evaluation of whether or not such Frankfurt-style cases, cases like Fischer’s given above, succeed in showing that PAP is false (i.e., succeed in meeting all three conditions of an FR scenario), an important limitation on condition (i) of an FR scenario must be discussed. When the proponent of PAP claims that having alternate possibilities to what one does is required for being morally responsible for so doing, the proponent of PAP is *not* claiming that having just any alternatives is what is necessary for moral responsibility. Consider a case in which Jones chooses at t_0 to rob a bank and, consequently, robs a bank at t_1 , and does so under no compulsion from any internal or external sources. It could always be claimed that, rather than choosing at t_0 to rob a bank, it was possible that a meteor fall on Jones at t_0 , or that Jones is teletransported by aliens to another universe at t_0 , or that Jones spontaneously combusts at t_0 , etc. Although it is possible that any of these things happen at t_0 as an alternative to what Jones actually does

at that time, none of these things would count as an alternate possibility that could plausibly be claimed relevant to explaining why Jones is morally responsible for what he actually does at t_0 , choose to rob a bank. More plausibly, the proponent of PAP is claiming that having only a certain kind of alternate possibility to what one does is what is necessary for being responsible for so doing. We call this certain kind of alternate possibility a robust alternate possibility, which is the sort of possibility that it would be reasonable or intuitive to claim that having this kind of alternate possibility is what (at least partially) explains why an agent is responsible for what he actually does; according to the proponent of PAP, an agent is morally responsible for doing something (at least partly) *in virtue* of having a robust alternate possibility (i.e., having a robust alternate possibility is explanatorily relevant to an agent's moral responsibility). A way to understand what the proponent of PAP means by 'robust alternative possibility' is to imagine a case in which all other conditions one believes necessary for moral responsibility are met (that is, all conditions other than having robust alternative possibilities; for example, one may believe that it is necessary that the agent does not have certain forms of mental illness that result in compulsion, that the agent is not under any form of constant direct brain manipulation by another agent, etc.), and then consider what kind of alternative(s) would be relevant to assessing whether or not the agent is morally responsible for what he actually does; those alternatives that would be relevant are robust alternatives and those that would not be relevant are not robust alternatives. According to the proponent of PAP, then, having a robust alternative possibility to what one does is required for being responsible for so doing.⁸

⁸ Thus, PAP should be reformulated as the following: having a *robust* alternative possibility to choosing/doing x is necessary for an agent to be morally responsible for choosing/doing x .

If, then, an FR scenario is meant to be a counterexample to PAP, where PAP is the claim that having robust alternative possibilities to what one does is required for being morally responsible for so doing, the first condition of an FR scenario must be changed to read: the agent has no *robust* alternate possibilities. So, in order for a case to be a counterexample to PAP, the following three conditions of an FR scenario must be satisfied: (1) the agent has no robust alternate possibilities to what he does, (2) whatever circumstances make it the case that the agent has no robust alternate possibilities are not what causes him to do what he does, or brings about what he does, and (3) the agent is morally responsible for what he does. Whether or not Frankfurt-style cases satisfy all three conditions of an FR scenario and, so, whether or not Frankfurt-style cases are counterexamples to PAP, will be the topic of the rest of this essay.

In the second section I will discuss the main line of objection to Frankfurt-type cases made by proponents of PAP, as well as the replies given by the proponents of Frankfurt-style cases. The proponent of PAP objects that Frankfurt-style cases are faced with a dilemma. I will discuss both horns of the dilemma, as well as the strategies used by the proponents of Frankfurt-style cases to respond to the dilemma, focusing most of my discussion on the second horn of the dilemma (which, I believe, is more interesting). Concerning the second horn of the dilemma, the proponents of Frankfurt-style cases have devised four strategies for finding a way out of the dilemma. I will discuss the first three strategies, showing that they are all problematic strategies to take as responses to the dilemma defense of PAP. The fourth strategy seems to me to be the strongest strategy and is taken by Derk Pereboom but, as I will show, still problematic. I will show that the success of Pereboom's case depends upon the correctness of his definition of

‘robustness.’ In the third section I will argue that Pereboom’s definition of ‘robustness’ is not correct, and I will argue for an alternative definition of ‘robustness,’ which if correct, would show that Pereboom’s Frankfurt-style case does include robust alternate possibilities and, so, does not satisfy condition (1) of an FR scenario. If, then, the definition of ‘robustness’ that I propose is correct, then Pereboom’s Frankfurt-style case will not satisfy all three conditions of an FR scenario and, so, will not be a counterexample to PAP.

AN OBJECTION TO FRANKFURT-STYLE CASES, AND REPLIES

There are many sorts of objections to Frankfurt-style cases given by defenders of PAP and in what follows I will be discussing the most prominent objection to Frankfurt-style cases, the objection that has been given the most attention in the recent literature, and which I take to be the most serious objection to Frankfurt-style cases. This objection is called the dilemma defense of PAP¹, initially suggested by Robert Kane² and developed by David Widerker³.

The dilemma defense of PAP begins by first pointing out that a standard Frankfurt-style case, such as Fischer's case given in the previous section, does not specify whether or not the occurrence of the involuntary prior sign is sufficient for the agent to make the choice he actually makes, *C*. Recall that in a standard Frankfurt-style case, an agent makes some choice, *C*, and that some device is in place to prevent the agent from choosing otherwise. The device does its job of preventing the agent from choosing otherwise in the following way: if the agent will make choice *C*, he will show some involuntary sign prior to making choice *C* (and with the occurrence of this prior sign, the device will not intervene, will remain idle, and will not affect what the agent does), such that the absence of this sign is the cue for the device to intervene and bring about choice *C* in the agent. In the actual scenario the involuntary prior sign is displayed, the device

¹ Widerker and McKenna, 2003, p. 8.

² Kane, 1985, p. 51 and, Kane, 1996, pp. 142-145.

³ Widerker, 1995, pp. 247-261.

does nothing and the agent makes choice C , but if this sign had not been displayed, the device would have intervened and brought about choice C (i.e., the absence of the prior sign is sufficient for the device to intervene and bring about choice C). Although the absence of the prior sign is sufficient for the device to intervene, we can see that it is not specified whether or not the occurrence of the involuntary prior sign in the actual scenario is sufficient for the agent to make choice C .

Suppose that it is stipulated that the occurrence of the involuntary prior sign in a standard Frankfurt-style case is sufficient for the agent to make choice C^4 , the choice the agent actually makes without the intervention of the device. If this is stipulated, then the occurrence of the involuntary prior sign determines that the agent will make choice C . If the agent's choice is determined by the occurrence of the involuntary prior sign, then it is not so obvious that the standard Frankfurt-style case meets the third condition of an FR scenario, that the agent is morally responsible for his actual choice. If an agent's choice is determined by the occurrence of an involuntary prior sign, it is not clear that we should claim that the agent is morally responsible for making such a choice, or for performing the action that is chosen, because it is not clear that we should hold people responsible for choices and actions that are the product of something involuntary, something the person had absolutely no control over. Aside from it not being clear that the agent is morally responsible for his actual choice if it is stipulated that the occurrence of the involuntary prior sign is sufficient for the actual choice made by the agent in a standard Frankfurt-style case - in which case a standard Frankfurt-style case would not succeed in showing

⁴ Hereafter, the suggestion that the occurrence of the involuntary prior sign is sufficient for the agent to make choice C is to be taken to mean either (a) the occurrence of the involuntary prior sign is sufficient for the agent to make choice C , or (b) the condition(s) that produces the involuntary prior sign is (are) sufficient for the agent to make choice C .

PAP false because such cases would not meet the third condition of an FR scenario - to claim that the agent is morally responsible under such circumstances would be to beg the question against the Incompatibilist who claims that we cannot be morally responsible for choices and actions that are determined. So, if we assume that the occurrence of the involuntary prior sign is sufficient for the agent's actual choice in a standard Frankfurt-style case, then (a) it is not clear that a standard Frankfurt-style case shows PAP false because it is not clear that the third condition of an FR scenario, that the agent is morally responsible for his actual choice, is met by such cases, and (b) to claim that the third condition is met by such cases would beg the question against the Incompatibilist.

That (a) and (b) result from stipulating that the occurrence of the involuntary prior sign is sufficient for the agent's actual choice in a standard Frankfurt-style case is problematic for the success of such cases as counterexamples to PAP under this stipulation. Concerning (a), part of the appeal of Frankfurt-style cases is that the intuition that the agent is morally responsible for the actual choice he makes is a very plausible intuition. Since in the actual scenario of a standard Frankfurt-style case the agent makes his choice as we believe choices are normally made by adults – free from compulsion, constraint, intervention by any device, etc. – the intuition that he is morally responsible is very reasonable. Part of the appeal of Frankfurt-style cases as counterexamples to PAP, where PAP is itself also an intuitively plausible principle, rests on the high degree of reasonability of claiming that the agent is morally responsible for his actual choice even though he could not have chosen otherwise. So, in so far as the success of Frankfurt-style cases as counterexamples to PAP relies on the high degree of reasonability of claiming that the agent is morally responsible for his actual choice, given that (a) results from the

stipulation that the occurrence of the involuntary prior sign is sufficient for the agent's actual choice in standard Frankfurt-style cases, stipulating this is problematic for the success of such cases as counterexamples to PAP.

Concerning (b), if the stipulation that the occurrence of the involuntary prior sign is sufficient for the agent's actual choice in a standard Frankfurt-style case means that the agent's actual choice was determined, then to claim that the agent is morally responsible for his choice would be to beg the question against the Incompatibilist. To see why this is a problem for the use of Frankfurt-style cases as counterexamples to PAP, recall that PAP is one premise of the Leeway Incompatibilist's argument for the incompatibility of moral responsibility and determinism: Having robust alternate possibilities to something one does is necessary for being morally responsible for so doing (PAP), but determinism removes alternate possibilities, and, so, moral responsibility and determinism are incompatible. If using a standard Frankfurt-style case as a counterexample to PAP is to show that the Leeway Incompatibilist's argument fails *because Frankfurt-style cases show that PAP, one premise of this argument, is false*, it would be problematic if the success of such cases was based upon the falsity of the Incompatibilist's conclusion. If the proponent of Frankfurt-style cases wishes to use such cases to show that the Leeway Incompatibilist's argument for the conclusion that moral responsibility and determinism are incompatible fails, then it cannot be claimed both that the agent's choice in such a case was determined and that the agent is morally responsible for this choice. To claim both that the agent's choice was determined and that he is morally responsible for this choice as a part of an argument against an argument that concludes that moral responsibility and determinism are incompatible would be to beg the question against the

Incompatibilist. We can see, then, that in so far as the proponent of Frankfurt-style cases wishes to use such cases to disprove a premise of the Leeway Incompatibilist's argument for the incompatibility of moral responsibility and determinism, the stipulation that the occurrence of the involuntary prior sign is sufficient for the agent's choice in Frankfurt-style cases is problematic for the success of such cases as counterexamples to PAP.

In the above given explanation of the first part, or first horn, of the dilemma defense of PAP, I have attempted to defend the following claim made by the proponents of the dilemma defense of PAP: if the involuntary prior sign determines the agent's choice in a standard Frankfurt-style case, then to claim that the agent is morally responsible for that choice begs the question against the Incompatibilist. While I do find it problematic for the proponent of Frankfurt-style cases to claim that the agent is morally responsible for his choice if this choice was determined by an involuntary prior sign, I do not believe that it is correct to claim that it begs the question against the Incompatibilist to do so. Considering that the Incompatibilist has his own arguments to support the claim that moral responsibility and determinism are not compatible, the Incompatibilist may not share the intuition that the agent in a Frankfurt-style case is morally responsible for his choice if this choice is determined. But, nevertheless, many will have this intuition and this intuition is plausible. Frankfurt-style cases are intended to be counterexamples to one premise (PAP) of the Leeway Incompatibilist's argument for the incompatibility of moral responsibility and determinism, where PAP has its support in intuition, just as the claim that the agent is morally responsible has its support in intuition.

Begging the question against one's opponent occurs when one uses the falsity of the conclusion of one's opponent to prove the truth of one's own conclusion, without

having any reason/support for using the falsity of the opponent's conclusion in one's own argument. But, the proponent of Frankfurt-style cases does have a reason for denying that moral responsibility and determinism are incompatible: there is a case, namely a standard Frankfurt-style case, in which it is plausible to claim that an agent is morally responsible for this choice, even if this choice was determined. Unless, then, we want to claim that counterexamples to a premise of an argument (here PAP) beg the question against the conclusion of that argument (here the Leeway Incompatibilist's conclusion), it does not seem as if Frankfurt-style cases beg the question against the Incompatibilist if we suppose that his choice was determined by an involuntary prior sign.

This, though, is not to say that the stipulation that the agent's choice is determined is not in any other way problematic for the proponent of Frankfurt-style cases. The stipulation that the agent is morally responsible for his choice if that choice is determined is problematic if the proponent of Frankfurt-style cases intends for a Frankfurt-style case to be used in an argument for Compatibilism in something like the following way:

1. It is possible to give a case in which an agent, Jones, chooses to act and acts in the way we believe normal adults to act - free from compulsion, constraint, mental illness, etc., while his choice is determined by some involuntary prior sign.
2. If Jones chooses to and acts in this normal way, he is morally responsible for so doing.
3. Suppose that there is an intervening device that would have made Jones choose as he in fact does had he shown some sign of choosing otherwise. Given the placement of this intervening device, Jones could not have done otherwise than he actually did (i.e., he has no alternate possibilities).
4. Therefore, since Jones is morally responsible for his choice and action despite having no alternate possibilities, alternate possibilities are not required for moral responsibility (i.e., PAP is false).
5. Determinism is merely one way of removing alternate possibilities, but (given 4) this removal of alternate possibilities is irrelevant to an agent's moral responsibility for what he did (and, there is no other feature of determinism that is relevant to an agent's moral responsibility).
6. Therefore, determinism and moral responsibility are compatible (i.e., Compatibilism is true).

The problem with this argument for the proponent of Frankfurt-style cases is that it makes the use of the special feature (the counterfactually intervening device that eliminates alternative possibilities) of Frankfurt-style cases unnecessary. Notice that the conclusion, that determinism and moral responsibility are compatible, follows from the first and second premise of the argument given above because the first premise stipulates that the agent's choice was determined, while the second premise states that the agent is morally responsible for his choice/action. This means that the special feature of a Frankfurt-style case, that feature being the counterfactual intervener that eliminates the agent's ability to do otherwise, is playing no role in supporting the Compatibilist conclusion and, so, the use of Frankfurt-style cases become superfluous in proving the Compatibilist conclusion. If, though, the proponent of Frankfurt-style cases claims that Frankfurt-style cases in some way advances the debate concerning moral responsibility (or plays some role in this debate), the special features of a Frankfurt-style case must be doing some work that an ordinary case of an agent choosing to perform and performing some action would not do. Thus, it is problematic for the Compatibilist to use Frankfurt-style cases to make the argument outlined above. We will see later that John Martin Fischer uses an argument similar to the one outlined above in order to defend Frankfurt-style cases from the first horn of the dilemma defense of PAP.

The dilemma defense of PAP continues by proposing that if it is stipulated in Frankfurt-style cases that the occurrence of the involuntary prior sign is not sufficient for the agent's choice, it is not the case that the agent lacks the ability to choose otherwise, and, so, the first condition of an FR scenario would not be satisfied. Under this stipulation, the occurrence of the involuntary prior sign is merely a reliable indicator that

the agent will make choice *C* and, so, it is possible that the involuntary prior sign be displayed and yet the agent make some choice other than *C* or make no choice at all. If the occurrence of the involuntary prior sign does not determine that the agent will make choice *C*, then, it seems that the agent will have the possibility of making a choice other than *C*, and, so, if the occurrence of the involuntary prior sign is not sufficient for the agent to make choice *C*, it is not the case that Frankfurt-style cases satisfy the first condition of an FR scenario, that an agent have no (robust) alternate possibilities to what he chooses.

It seems, then, that the success of Frankfurt-style cases as counterexamples to PAP is jeopardized by the stipulation that the occurrence of the involuntary prior sign is not sufficient for the agent to make choice *C*. But, as we saw in the previous discussion, the stipulation that the involuntary prior sign is sufficient for the agent to make choice *C* also jeopardized the success of Frankfurt-style cases as counterexamples to PAP. The proponent of Frankfurt-style cases is, then, posed with the following dilemma: If, on the one hand, it is stipulated that the occurrence of the involuntary prior sign is sufficient for the agent to make choice *C*, then the agent cannot (unproblematically) be claimed to be morally responsible for his choice, and, so, Frankfurt-style cases do not satisfy the third condition of an FR scenario. If, on the other hand, it is stipulated that the occurrence of the involuntary prior sign is not sufficient for the agent to make choice *C*, then the agent can choose otherwise and, so, Frankfurt-style cases do not satisfy the first condition of an FR scenario. Either way, it is argued by the proponent of the dilemma defense of PAP, standard Frankfurt-style cases cannot satisfy all three conditions of an FR scenario and, so, such cases are not counterexamples to PAP.

The First Horn of the Dilemma: Assumption that the Prior Sign Determines the Agent's Choice (i.e., That the Prior Sign Deterministically Causes the Agent's Choice)

One way that the proponent of Frankfurt-style cases can respond to the dilemma defense of PAP is to argue that the question is not begged against the Incompatibilist⁵ if it is assumed that the occurrence of the prior sign determines the agent's choice. The proponent of Frankfurt-style cases can argue his way out of the dilemma by arguing that to assume that the agent in a Frankfurt-style case is morally responsible for his choice given the stipulation that the occurrence of the prior sign is sufficient for (or, determines) the agent's choice is not to beg the question against the Incompatibilist. If the proponent of Frankfurt-style cases could successfully argue that doing so would not beg the question against the Incompatibilist, then the proponent of Frankfurt-style cases could claim both that the occurrence of the prior sign is sufficient for the agent's choice and that all three conditions of an FR scenario are satisfied, thus showing that he is not faced with a dilemma.

In "Frankfurt-type Examples and Semi-Compatibilism" John Martin Fischer argues that the use of standard Frankfurt-style cases in an argument against PAP need not be seen as begging the question against the Incompatibilist if it is assumed that the occurrence of the prior sign determines the agent's choice:

Begin with the first horn of the dilemma: the assumption that causal determinism obtains. I agree that one cannot now simply and precipitously conclude, from

⁵ While I have argued above that the problem with the assumption that the prior sign determines the agent's choice is not that the question is begged against the Incompatibilist but, rather, that the use of Frankfurt-style cases becomes unnecessary to establish the Compatibilist's conclusion, my main objective in these next few paragraphs is to explain the debate concerning this horn of the dilemma as it has played out in the recent literature as a debate between John Martin Fischer and Stuart Goetz. Both Fischer and Goetz have framed the discussion as a discussion of whether or not the question is begged against the Incompatibilist and, so, I will frame the debate as they do for the time being, and then make some critical remarks after explaining the discussion between Fischer and Goetz.

consideration of the examples, that the agent is morally responsible for his choice and behavior. But in any case this is not the way I would have proceeded; I never have envisaged a simple “one-step” argument to the conclusion that (say) Jones is morally responsible for his choice and action. Rather, I employ the Frankfurt-type examples as the first (but obviously important) step of a slightly more complex argument to the conclusion that Jones is morally responsible for his choice and action (despite lacking alternative possibilities).

The argument goes as follows. First, one carefully considers the Frankfurt-type cases. Upon reflection, I believe that one should conclude that in these cases the lack of alternative possibilities does not in itself ground a claim that the agent is not morally responsible for his choice and action. In other words, I think that the examples make highly plausible the preliminary conclusion that *if* Jones is not morally responsible for his choice and action, this is *not* simply because he lacks alternative possibilities. After all, everything that has any causal (or any other kind of) influence on Jones would be exactly the same, if we “subtracted” [the intervening device] from the scene. And Jones’s moral responsibility would seem supervenient on what has influence or impact on him in some way.

So the relevant (preliminary) conclusion is, if Jones is not morally responsible for his choice and action, the reason is not simply that he lacks alternative possibilities. And it does *not* appear to beg the question to come to this conclusion, even if causal determinism obtains. The first step is to argue, based on the Frankfurt-type examples, that intuitively it is plausible that alternative possibilities are irrelevant to ascriptions of moral responsibility . . . The *second* step in the argument consists in asking whether causal determinism *in itself and apart from ruling out alternative possibilities* threatens moral responsibility. I have considered various possible reasons why someone might think that causal determinism does threaten moral responsibility in itself and apart from ruling out alternative possibilities, and I have come to the conclusion that it is not plausible to accept any of these reasons. It seems to me that this two-stage argument is highly plausible and does not beg the question against the Incompatibilist, even on the assumption of causal determinism. Thus I believe that the use of the “prior sign” cases can be defended against the charge of begging the question.⁶

I have given this lengthy quote from Fischer’s essay because this argument is the starting point of Stewart Goetz’s criticism of defending Frankfurt-style cases from the charge of begging the question. Before explaining Goetz’s criticism it is important to get clear on what Fischer’s two-step argument is. I take Fischer’s two-step argument against

⁶ Fischer, 2002, pp. 291-292.

PAP to be the following, where 7-9 constitutes Fischer's first step and 10-12 constitutes his second step:

7. An agent's moral responsibility, or lack thereof, for some choice and/or action can only be based upon those things which exerted influence over the agent's making this choice or performing this action.
8. In the actual scenario of a Frankfurt-style case, that the agent has no alternative possibilities does not influence the agent's actual choice and action, because this lack of alternative possibilities results from the intervening device, which does not exert any influence over the agent's actual choice and action.
9. Therefore, in a Frankfurt-style case, the agent's lack of alternative possibilities cannot be the basis for the claim that the agent is not morally responsible for his choice and action. (i.e., lack of alternative possibilities is not sufficient for not being morally responsible, i.e., PAP is false, i.e., alternative possibilities are irrelevant to ascriptions of moral responsibility).
10. The agent in a Frankfurt-style case is not morally responsible for his determined choice only if either (c) alternative possibilities are relevant to ascriptions of moral responsibility, or (d) determinism directly, and not by ruling out alternative possibilities, undermines moral responsibility.
11. I (Fischer) have considered various reasons for believing (d) and have come to the conclusion that (d) is false, and it has been shown in 3 above that (c) is false.
12. Therefore, the agent in Frankfurt-style cases is morally responsible for what he does, even if his choice is determined by the occurrence of the prior sign.
In his "Frankfurt-style Counterexamples and Begging the Question," Stewart Goetz

replies to Frankfurt's two-step argument by arguing that this two-step argument itself begs the question against the Incompatibilist by claiming that the agent is morally responsible for his choice while assuming determinism in the actual sequence. The heart of Goetz's criticism of Fischer's two-step argument can be seen in the following three passages:

What is one to think of Fischer's two-step argument? In the end it is difficult to see how it advances the debate between Compatibilists and Incompatibilists. This is because contrary to what Fischer claims, his two-step argument begs the question against the Incompatibilist in the same way that the one-step arguments do: it assumes, because it requires, the truth of causal determinism in the actual sequence of events. It requires the truth of causal determinism to create the *illusion* that it is the presence of something in the alternative sequence of events (e.g., Black's device) that makes it the case that Jones is not free to choose otherwise. It is only through the creation of this illusion and the fact that Black's device is not explaining Jones' actual choice that one is tempted or inclined to endorse the

conclusion of the first step of Fischer's argument, which is that the lack of alternative possibilities is not sufficient for the lack of moral responsibility (PAP is false) and, thereby, that Jones might be morally responsible even though he is not free to choose otherwise. Once this illusion is exposed, one's initial conviction that the lack of alternative choice is sufficient for the lack of moral responsibility (PAP is true) is vindicated.

It is true that the advocate of PAP believes that the lack of the freedom to choose otherwise *is sufficient for* the lack of moral responsibility, and in that sense the latter obtains simply because of the obtaining of the former. It is incorrect, however, to assume that a person who affirms PAP believes that the lack of moral responsibility obtains simply because of the obtaining of the freedom to choose otherwise in the sense that the obtaining of the latter *by itself*, that is, apart from what explain it, *explains* the obtaining of the former. The proponent of PAP thinks that the lack of the freedom to choose otherwise does not by itself explain the absence of moral responsibility. This is because he believes that when this lack obtains, its obtaining is itself explained by, and only can be explained by, the occurrence of causal determinism in the actual sequence of events.

When an agent is not free to choose otherwise, not only does causal determinism explain what occurs in the actual sequence of events, but also it explains the lack of the freedom to choose otherwise. Frankfurt himself recognized this *prima facie* dual explanatory role of causal determinism in his classic discussion of PAP. Thus, he states that "[i]n seeking illustrations of the principle of alternative possibilities, *it is most natural to think* of situations in which the same circumstance both bring it about that a person does something [chooses] and make it impossible for him to avoid doing it [choosing otherwise]"⁷ Implicit in Frankfurt's remark is the idea that it is intuitively plausible to think not only that causal determinism in the actual sequence of events *is capable of* producing situations wherein the same circumstances produce the actual choice and make it impossible for the agent to choose otherwise, but also that causal determinism in the actual sequence of events is the *only* way to bring about situations in which an agent does not have the possibility of choosing otherwise. It is because it is intuitively plausible to think this that Frankfurt *assumes the burden* of constructing examples wherein the circumstances that make it impossible for Jones to choose otherwise are not the same as or identical with what is going on in the actual sequence of events. The reason why Frankfurt's paper is so controversial is that it claims to provide an example where an agent does not have an alternative choice and the lack of this alternative is *not explained* by causal determinism in the actual sequence. Until such an example is provided, the incompatibilist is on firm ground in believing that the lack of the freedom to choose otherwise is explained by and can only be

⁷ Frankfurt, 1969.

explained by causal determinism in the actual sequence of events and that PAP is true.⁸

Essentially, Goetz is arguing that premise 8 (the second premise) of Fischer's two-step argument against PAP is false, and that the initial plausibility of this premise is due to the following illusion: that what suffices for the agent's lack of alternative possibilities in a Frankfurt-style case is only effective (or, active) in the counterfactual scenario (i.e., does not exert any influence in the actual scenario). Goetz argues that the truth of premise 8 above relies on the illusion that what explains, or suffices for, the agent's lack of alternative possibilities is the intervening device, which does not exert any influence over the agent's actual choice and action (this is the point of the first passage above), when what really and only can explain, or suffice for, the agent's lack of alternative possibilities is the obtaining of determinism (this is the point of the second and third passages above), which does influence – by necessitating - the agent's actual choice and action.

While Goetz is, mistakenly I believe (for the reasons laid out in the earlier discussion of this first horn of the dilemma), ultimately trying to argue that Fischer's two-step argument begs the question against the Incompatibilist and, so, does not help the proponent of Frankfurt-style cases respond to the first horn of the dilemma defense of PAP, I believe that Goetz is pointing to something that is problematic with Fischer's defense of Frankfurt-style cases. So, even though Goetz mistakenly charges Fischer's two-step argument with begging the question against the Incompatibilist, Goetz is correct to point out that premise 8 of Fischer's two-step argument relies on an illusion and, so, is

⁸ Goetz, 2005, pp. 87-89.

unsupported. A look at Fischer's reply to Goetz's criticism will bring out this mistake more clearly. Fischer replies that:

In the Frankfurt-type scenarios, *two* causes make it the case that Jones is unable to choose otherwise at T_2 : the prior condition of the world (together with the laws of nature) and Black's counterfactual intervention. What the examples show is that *the mere fact* that Jones is unable to choose otherwise does not *in itself* establish that Jones is not morally responsible for his choice. This is because Black's counterfactual intervention is one of the factors that make it the case that Jones is unable to choose otherwise at T_2 , and yet it is irrelevant to the grounding of Jones's moral responsibility. Considering this factor (the counterfactual intervention), and bracketing any other factor that might make it the case that Jones is unable to choose otherwise at T_2 , it seems to me that Jones may well be morally responsible for his action. The mere fact that he lacks alternative possibilities, then, cannot in itself be the reason Jones is not morally responsible, if indeed he is not morally responsible.

Now, of course, it is *also* true that the prior condition of the world, together with the natural laws, make it the case that Jones lacks alternative possibilities. But, given that the mere fact of lacking alternative possibilities does not in itself rule out moral responsibility, why should *this way* of lacking alternative possibilities rule out moral responsibility? Why exactly should the significance of causal determinism be that it rules out alternative possibilities? This is exactly what the Frankfurt-type examples call into question.⁹

I take it that Fischer's reply to Goetz's argument amounts to the following: the agent's lack of alternative possibilities in a Frankfurt-style case is brought about by two things, one of which is the intervening device that does not influence the agent's actual choice and action and, so, is irrelevant to the agent's moral responsibility. Given that one way of removing the agent's alternative possibilities is irrelevant to the agent's moral responsibility, there seems no reason to believe that any other particular way of removing alternative possibilities is relevant to an agent's moral responsibility.

Therefore, the obtaining of determinism, which is merely one other way of removing alternative possibilities, is irrelevant to an agent's moral responsibility and, so, it would

⁹ Fischer, 2006, pp. 340.

not beg the question against the Incompatibilist to claim that the agent in a Frankfurt-style case is morally responsible if his choice was determined. (Notice the similarity between this argument and the argument discussed earlier, the argument I claimed to be problematic for the Compatibilist).

Fischer's reply to Goetz is mistaken, and the mistake lies in his first claim: "In the Frankfurt-type scenarios, *two* causes make it the case that Jones is unable to choose otherwise at T_2 : the prior condition of the world (together with the laws of nature) and Black's counterfactual intervention."¹⁰ When Fischer claims that "two causes make it the case that Jones is unable to choose otherwise,"¹¹ he is claiming that there are two causes that are individually sufficient for Jones being unable to choose otherwise: the obtaining of determinism and the counterfactual intervener. Fischer must be claiming that the obtaining of determinism and the intervening device are each sufficient to make it the case that Jones cannot choose otherwise, because if he is not claiming this, premise 9 of his two-step argument (that the lack of alternative possibilities is irrelevant to ascriptions of moral responsibility) would not follow from premise 7 (only factors which exert influence over what an agent does are relevant to ascriptions of moral responsibility) and premise 8 (the presence of the intervening device does not exert any influence over the agent's choice in a Frankfurt-style case). If he is not relying on the claim that the intervening device is sufficient for the agent's lack of alternative possibilities, then it would still be possible that there is something that is both sufficient for it and does exert influence over the agent's choice and, so, the conclusion drawn in premise 9 would not

¹⁰ Fischer, 2006, pp. 340.

¹¹ Fischer, 2006, pp. 340.

follow from premises 7 and 8. In order for his two-step argument to be sound, then, when Fischer claims that “two causes make it the case that Jones is unable to choose otherwise,”¹² we must take him to be claiming that the presence of the counterfactual intervener is alone sufficient for Jones’s inability to choose otherwise.

It is not correct, though, to claim that the intervening device is sufficient for the agent’s lack of alternative possibilities in a Frankfurt-style case. The intervener cannot do its job of being a counterfactual intervener unless it is the case that the occurrence of the involuntary prior sign determines that the agent will make choice *C* and, so, the presence of the device is not sufficient to remove the agent’s ability to do otherwise. It must also be the case that a deterministic relationship exists between the occurrence of the prior sign and the agent’s making choice *C*. If this deterministic relationship is not present between the occurrence of the prior sign and the agent’s making choice *C*, then it could be the case that the prior sign is displayed (in which case the device does not intervene) and yet the agent makes a choice other than *C*. So, Fischer is not correct to claim that the presence of the intervening device is sufficient for the agent’s lack of alternative possibilities, because it must also be the case that the occurrence of the prior sign determines that the agent will make choice *C*. This is the illusion that Fischer is relying on in making his two-step argument, namely, that the cause of the agent’s lack of alternative possibilities has no (causal) influence over what the agent chooses; but, surely that the occurrence of the prior sign determines that the agent will make choice *C* does have (causal) influence over what the agent chooses. We can see, then, that although

¹² Fischer, 2006, pp. 340.

Goetz mistakenly labels Fischer's two-step argument as begging the question against the Incompatibilist, Goetz has pointed to a serious mistake in Fischer's two-step argument.

According to Goetz, Fischer's two-step argument does not give the proponent of Frankfurt-style cases a satisfactory response to the dilemma defense of PAP. Goetz gives the following summary of the debate concerning the success of Frankfurt-style cases as counterexamples to PAP, and presents the following challenge to the proponent of Frankfurt-style cases:

So, where are we? Right back where we began, because the dilemma for proponents of FSCs is still this: either causal determinism obtains in the actual sequence, in which case Black's device is not causally efficacious in preventing Jones from choosing otherwise, or causal determinism fails to obtain in the actual sequence, in which case Black's device is also not causally efficacious in preventing Jones from choosing otherwise. If my foregoing reasoning is correct, it is still the case that Fischer has not provided any satisfactory Compatibilist way through the horns of this dilemma and, as a result, PAP is not undermined. To find a way through the horns of this dilemma and undermine the intuitively powerful that PAP captures, it remains the case that Fischer (any Frankfurtian) must provide an example where what explains the agent's lack of the freedom to choose otherwise is something other than causal determinism in the actual sequence of events.¹³

According to Goetz, then, in order for a Frankfurt-style case to be successful in showing that PAP is false, it must be the case that the prior sign is not sufficient for the agent's actual choice (i.e., that the occurrence of the prior sign does not determine that the agent will make choice *C*), because if the prior sign is sufficient for the agent's choice, then this sufficiency is what explains the agent's lack of alternative possibilities, and this sufficiency is causally relevant to what the agent chooses (in which case the second condition of a FR scenario would not be satisfied).

¹³ Goetz, 2005, p. 94.

Given the above discussion, it seems to me that an attempt to rescue Frankfurt-style cases from the dilemma defense of PAP by responding to the first horn of the dilemma is not a promising way for the proponent of Frankfurt-style cases to respond to the dilemma defense of PAP. As an alternative to responding in this way, the proponent of Frankfurt-style cases can try to show that Frankfurt-style cases can be rescued from the dilemma defense of PAP by giving a case in which the prior sign does not determine the agent's choice (or, in which there is no prior sign) and in which the agent cannot choose/do otherwise and is morally responsible for so choosing/doing.

The Second Horn of the Dilemma: Assumption that the Prior Sign is Not Sufficient for the Agent's Choice (i.e., That the Agent's Choice Is Indeterministically Caused)

Recall that the second horn of the dilemma defense of PAP brings the following problem to the table: if it is stipulated that the occurrence of the prior sign is not sufficient for the agent's choice and, so, the agent's choice is not determined by the prior sign (i.e., if the agent's choice is brought about indeterministically), then it seems that the prior sign is merely a reliable indicator that the agent will make choice some choice, *C*. If the occurrence of the prior sign is merely a reliable indicator that the agent will make choice *C*, it would not follow that the agent could not choose otherwise than *C* if the prior sign occurs. So, the problem for the proponent of Frankfurt-style cases is that if the prior sign is not sufficient for the agent's actual choice, the first condition of an FR scenario (that the agent cannot choose otherwise, that the agent has no robust alternative possibilities) cannot be met, and, so, a Frankfurt-style case cannot show that PAP is false.

Many proponents of Frankfurt-style cases, though, have attempted to rescue Frankfurt-style cases from the dilemma defense of PAP by showing that a case can be given in which both the agent's choice is brought about indeterministically and in which

the first condition of an FR scenario is met – the agent cannot choose otherwise, or, does not have any robust alternate possibilities. In their introduction to *Moral Responsibility and Alternative Possibilities*, editors David Widerker and Michael McKenna give brief summaries of the four strategies that have developed to give such cases, cases in which both the agent's choice is brought about indeterministically and in which the agent cannot choose otherwise, or has no robust alternative possibilities. In what follows I will explain each of these four strategies, giving an example of a case that uses each strategy. I believe that the fourth strategy to be discussed, the strategy taken by Pereboom, is the strongest strategy, and so I will devote most of the discussion that follows to evaluating Pereboom's Frankfurt-style case. I will, though, briefly discuss the problems I see with the first three strategies as well.

No-Prior-Sign Cases

Those proponents of Frankfurt-style cases that give No-Prior-Sign cases try to develop a situation in which an agent cannot choose otherwise, and this is ensured without the use of a prior sign as a basis for intervention. In No-Prior-Sign cases, the agent cannot choose otherwise because there are two processes that proceed independently of one another and which both end in the agent's making some choice, *C*. One process, the process that actually results in the agent (freely) making choice *C*, is an indeterministic process, and the other process is a deterministic process that will bring about choice *C* in all cases except the case in which the other process, the indeterministic process, brings about this choice. Since the agent makes choice *C* as a result of an indeterministic process, we can hold the agent morally responsible for his choice and, yet, the agent could not do otherwise than make choice *C*. A No-Prior-Sign case is given by Alfred Mele and David Robb:

At t_1 , Black initiates a certain deterministic process P in Bob's brain with the intention of thereby causing Bob to decide at t_2 (an hour later, say) to steal Ann's car. The process, which is screened off from Bob's consciousness, will deterministically culminate in Bob's deciding at t_2 to steal Ann's car unless he decides on his own at t_2 to steal it or is incapable at t_2 of making a decision (because, for example, he is dead at t_2). (Black is unaware that it is open to Bob to decide on his own at t_2 to steal the car; he is confident that P will cause Bob to decide as he wants Bob to decide.) The process is in no way sensitive to any "sign" of what Bob will decide. As it happens, at t_2 Bob decides on his own to steal the car, on the basis of his own indeterministic deliberation about whether to steal it, and his decision has no deterministic cause. But if he had not just then decided on his own to steal it, P would have deterministically issued, at t_2 , in his deciding to steal it. Rest assured that P in no way influences the indeterministic decision-making process that actually issues in Bob's decision.¹⁴

If this case is plausible, it seems to meet all three conditions of an FR scenario: (1) the agent cannot choose otherwise at t_2 than to steal Ann's car, (2) what brings about the circumstance that the agent cannot choose otherwise – the deterministic process P – in no way influences the agent's actual choice, and (3) the agent is morally responsible for his actual choice to steal Ann's car, because he did so on his own through an indeterministic decision-making process. I do not, though, believe that this case can meet all three conditions of an FR scenario given that it is faced with the following dilemma: in terms of processes x (the indeterministic process that results in the agent's choice to steal the car) and P (the deterministic process), if we understand each process as being a temporal process with a beginning and an end, where at the instant the process ends, the choice made by the agent, the processes can either be such that:

- (e) process P is faster than process x (i.e., the deterministic process will end before the indeterministic process and, so, will result in the agent's choice at t_2 , before the indeterministic process has reached its end), or
- (f) both processes can be simultaneous (i.e., both processes would simultaneously end and, so, simultaneously produce the choice at t_2 , in which case the choice is overdetermined), or

¹⁴ Mele and Robb, 1998, pp. 101-102.

(g) process x can be faster than process P (i.e., the indeterministic process will end before the deterministic process and, so, result in the agent's choice at t_2 , before the deterministic process has reached its end).

In neither situation (e), (f), or (g) will all three conditions of an FR scenario be satisfied.

Suppose that (e) is the case and, so, the deterministic process results in the agent's choice at t_2 to steal the car. If this is the case, then (2) above is not satisfied (and it would beg the question to claim that (3) is satisfied) and, so, Mele and Robb's case would not be an FR scenario. Suppose, next, that (f) is the case, and the two processes are simultaneous so that the agent's choice at t_2 is overdetermined. It then seems unclear whether (2) above is satisfied, and this lack of clarity, I believe, results from it being not clear whether the existence of a deterministic process usurps any other process as the (and the only) causal explanation of the end result of such a process. If both a deterministic and an indeterministic process simultaneously result in the very same thing, it is hard to see what role the non-deterministic process is playing, since the deterministic process alone is sufficient for, while the non-deterministic process is not sufficient for, the end in question. Even though (f) stipulates that both the deterministic and indeterministic processes end at the same time and, so, both cause the end in question, it is not the case that the indeterministic process is sufficient for the end in question, while it is the case that the deterministic process is sufficient for the end in question. A deterministic process that has some result x at time t_n is one which, given the past and the laws of nature, could not have had any result other than x at t_n , whereas an indeterministic process that has result x at time t_n is one which, given the past and the laws of nature, could have had a result other than x at t_n ; indeterministic processes, then, are not sufficient for their ends or results. It is not clear, then, that if the two processes occur simultaneously, that the deterministic process is not influencing the agent's choice at t_2 , because it seems

plausible to claim that if the two processes simultaneously result in the choice, then the deterministic process has (at least, and perhaps all there is to it) something to do with the agent's making that choice.¹⁵ Suppose, finally, that (g) is the case and the indeterministic

¹⁵ Mele and Robb address this in their paper, and I take it that they intend that (f) is the case, that the two processes occur simultaneously and yet the indeterministic process is THE process that results in the agent's choice at t2 to steal the car: "The coherence of our scenario may, however, be called into question. How, one might wonder, can it happen that Bob decides on his own at t2 to steal Ann's car, given the presence of the deterministic process we mentioned? One can understand how, prior to t2, Bob might decide on his own to steal the car. (Notice that in that case, other things being equal, Bob could have decided otherwise at this earlier time . . .). But how can it happen that Bob decides on his own at t2 to steal the car, and that P does not produce the decision, given what we said about P?"

Consideration of the following fanciful machine will prove useful in answering this question. The machine, designed by a specialist in machine art, produces artistic widgets of different shapes and colors. The colors of the widgets produced are determined by the color of a ball bearing (bb) that hits the machine's receptor at a relevant time. The machine, M, is surrounded by several automatic bb guns, each containing bbs of various colors. The relevant aspect of M's mechanical design, for our purposes, is relatively simple. First, with one qualification, if a bb of color x hits M's receptor, and M is not already in the process of making a widget, M at once starts a process designed to result in the production of an x-colored widget. Second, because two or more bbs sometimes hit the receptor simultaneously, the artist has designed his machine in such a way that whenever this happens (while M is not busy making a widget) M at once starts a process designed to result in the production of a widget the color of the right-most bb.

Bob is analogous to M in an important respect. He is physically and psychologically so constituted that if an unconscious deterministic process in his brain and an indeterministic decision-making process of his were to "coincide" at the moment of decision, he would indeterministically decide on his own and the deterministic process would have no effect on his decision. This situation is an analogue of a case in which two bbs of the same color simultaneously hit M's receptor (while M is not busy making a widget)" [Mele and Robb, 1998, pp. 103-104].

It seems to me that the consideration of this fanciful widget-making machine does not help Mele and Robb's case to be an FR scenario. Recall that Mele and Robb's case is intended to be a case in which the agent could not have chosen otherwise than to steal Ann's car, whatever makes it the case that the agent could not choose otherwise does not exert influence over what the agent chooses, and the agent is morally responsible for his choice because he made it indeterministically. Stipulating that the agent makes the choice to steal Ann's car indeterministically is what allows Mele and Robb's case to be a case that satisfies the second and third of these conditions despite the fact that a deterministic process makes it the case that the agent could not have chosen otherwise. One should note, though, that merely stipulating that the agent's choice was made indeterministically does not amount to the stipulation that the agent's choice was made indeterministically in the right sort of way, where the right sort of way is the way such that the agent can be claimed morally responsible (at least partly) in virtue of the choice arising in this indeterministic fashion. In order that the agent's choice being brought about indeterministically is the reason, or one of the reasons, for which we claim that the agent is morally responsible for this choice, it must be the case that the indeterminacy is located in a proper place in the causal history of the choice. For example, if the indeterminacy was located at some point in the causal history of the choice that occurred billions of years ago, we would not claim that this is the right sort of indeterminacy to allow for the agent's moral responsibility for his choice billions of years later. Although it is very difficult to determine at what point in the causal history of a choice the indeterminacy must be in order that the agent is morally responsible for this choice, it seems most reasonable to claim that the indeterminacy must occur at the very moment the choice is made. It seems to me that when we claim that some choice was made indeterministically, this does not mean that something else happened indeterministically from which followed a choice; it seems to me that when we claim that some agent is morally responsible for some choice (at least partly) in virtue of the fact that the choice was made indeterministically, we do not mean to claim that some agent is morally

process results in the agent's choice at t_2 because it reaches its end before the deterministic process reaches its end. In this case, the indeterministic process will result in the agent's choice and, so, it is possible that the agent will make some choice at t_2 other than to steal the car. If (g) is the case, then, (1) above is not true and, so, Mele and Robb's case would not be an FR scenario. Either way, then, whether (e), (f), or (g) is the case, Mele and Robb's Frankfurt-style case does not meet all three conditions of an FR scenario.¹⁶

Blockage Cases

Blockage Examples also attempt to meet all the conditions of an FR scenario without the use of a prior sign. In a blockage case, all neural pathways other than the pathway that is (coincidentally) actually taken are blocked, so that (1) the agent could not

responsible for his choice (at least partly) in virtue of the fact that something happened indeterministically from which followed the agent's choice. The claim that the agent is morally responsible for some choice made only makes sense to me if it is the occurrence of the choice itself that is indeterministic; if it is not the case that up until the choice was made it could have been different than it was (i.e., if it is not the case that the occurrence of the choice is itself indeterministic), it is hard to see the relevance of the claim that an agent made some choice indeterministically to the claim that an agent is morally responsible for that choice.

In order, then, for the agent to be morally responsible for his choice at t_n , it must be the case that everything in the world could have been just as it was up to t_n and yet it was possible that the agent makes some other choice at that time. I would like to suggest that, given the setup of their case, the indeterministic process x that results in the agent's choice to steal Ann's car does not meet this condition. Taking into account the analogy with the widget-making machine, and given that (f) is the case and the deterministic process and the indeterministic process both hit the decision-making node(s) simultaneously, it must be the case that the processes hitting the decision-making node is a distinct event from the event of the choice being made in the same way that two bbs of the same color hitting the right and left receptors is a distinct event from the initiation of or the making of the widget. Since the indeterministic process hitting the decision-making node is a separate event from the event of the choice being made, in order for the choice to be made indeterministically in the right sort of way to allow for the agent to be morally responsible for this choice, it must be the case that everything prior (including the indeterministic process hitting the decision-making node) could have been the same as it was, and yet the agent could have chosen otherwise. In Mele and Robb's case, though, this is not so; once the indeterministic process hits the decision-making node, the agent will choose to steal Ann's car. And, so, I believe that we should not claim that the agent's decision in Mele and Robb's case is brought about indeterministically in the right sort of way to allow for the agent to be morally responsible for his choice and, so, Mele and Robb's case does not satisfy the conditions of an FR scenario and, so, is not a counterexample to PAP.

¹⁶ It should be noted that others have criticized Mele and Robb's case as well: See Pereboom, 2001, pp. 14-15; See also Kane, 2003; See also Widerker, 2000, pp. 183-185.

have done anything other than he did, but (2) the blockage is such that it in no way affects what the agent actually does and, so, (3) the agent is morally responsible for what he does. In “Moral Responsibility and Unavoidable Action,” David Hunt suggests such a blockage case that is based upon Locke’s locked door example:

The original “Frankfurt scenario” is of course John Locke’s famous example in which “a man be carried whilst fast asleep into a room where is a person he longs to see and speak with, and be there locked fast in, beyond his power to get out; he awakes and is glad to find himself in so desirable company, which he stays willingly in, i.e. prefers his stay to going away.”¹⁷ In this case the man is morally responsible for what he does even though he cannot do otherwise . . . But the condition that makes for unavoidability in this example is not counterfactual in nature: the door is actually locked; it doesn’t lock only when someone approaches the door and tries to leave. What makes the locked door compatible with the man’s moral responsibility is simply that it is not among the conditions actually leading the man to stay in the room . . . Locke’s example does provide some initial encouragement that the unavoidability essential to a Frankfurt scenario does not have to rest on a counterfactual device.

Thus encouraged . . . Imagine then a device that blocks neural pathways rather than doorways . . . The mechanism is not intervening directly in the series itself; it is allowing the series to unfold on its own, but simply blocking all alternatives to the series. Of course it can’t block alternatives *in response* to the way the series is unfolding, because then the blockage would be coming too late to have any effect on the avoidability or unavoidability of Jones’s actions. Instead the mechanism blocks alternatives in advance, but owing to a fantastic coincidence the pathways it blocks just happen to be all the ones that will be unactualized in any case, while the single pathway that remains unblocked is precisely the route the man’s thoughts would be following anyways (if all neural pathways were unblocked). Under these conditions, the man appears to remain responsible for his thoughts and actions . . .

¹⁸

It is clear that Hunt takes his case to satisfy all three conditions of an FR scenario: (1) the agent cannot do otherwise than he does (this is ensured by the mechanism, which, coincidentally, blocks all unactualized pathways before the actual pathway is actualized),

¹⁷ Hunt’s footnote reads: John Locke, *An Essay Concerning Human Understanding* Bk.II, ch. XXI, § 10. Of course Locke has his own agenda here, which is not entirely the same as Frankfurt’s.

¹⁸ Hunt, 2000, pp. 217-218.

(2) what brings about the agent's inability to do otherwise in no way influences what he does (because the operation of the mechanism does not affect the actual pathway), and (3) the agent is morally responsible for what he does. Robert Kane has challenged the intuition elicited by Hunt's blockage case, and Pereboom elaborates upon Kane's challenge, arguing that our intuition that (3) the agent is morally responsible in this case is not clearly correct, because it is far from clear whether it is reasonable to claim, given the circumstances, that the agent's choice was brought about indeterministically. Kane's challenge to such blockage cases is the following:

In [a case in which every other alternative is blocked except the agent's making A at t], of course there *are* no alternative possibilities left to the agent; every one is blocked except the agent's choosing A at t. But now we seem to have determinism pure and simple. By implanting the mechanism in this fashion, a controller would have predetermined exactly what the agent would do (and when); and, as a consequence, the controller, not the agent, would be ultimately responsible for the outcome. Blockage by a controller that rules out alternative possibilities is simply predestination; and on my view at least, predestination runs afoul of ultimate responsibility.¹⁹

Kane's claim here is that given the method of the mechanism, even though it only coincidentally blocks the pathways that are unactualized, since the mechanism does this before the actualized pathway is actualized, the blockage amounts to predestination of the agent's choice. If an agent's choice being predestined means, as it means according to Kane, that the agent is not ultimately responsible for such a choice, then, the presence of the blockage removes the agent's responsibility. Kane's objection to blockage cases, then, is that blocking alternative pathways prior to the actualization of the pathway that leads to the agent's choice amounts to determining the agent's choice, in which case, the

¹⁹ Kane, 2000, p. 162.

third condition of an FR scenario would not be met by such blockage cases (without begging the question).

In *Living Without Free Will*, Pereboom strengthens this criticism by first suggesting that the plausibility of claiming that the agent in a blockage case is morally responsible is due to the fact that the presence of the blockage does not have any part in the actual causal history of the agent's choice – it is purely coincidental that just those pathways were blocked that were unactualized (and would have been unactualized even if the blockages were not present). Consider, though, the following two situations: In situation A the agent makes choice x with no blockages or anything else unusual present (and so could have at that time chosen differently), while in situation B the agent makes choice x with the same causal history as in situation A, but in situation B the blockages are in place in the way that Hunt suggests – ahead of time such that coincidentally those pathways other than the actualized pathway are blocked.²⁰ Concerning situations A and B, Pereboom says that “cases of the above sort might be misleading just because it is natural to assume that the actual causal history of an event is essentially the same in each, given that the only difference between them is a restriction that would seem to have no effect

²⁰ I have paraphrased the relevant aspects of Situation A (the first case) and Situation B (the second case), the cases Pereboom uses at this point of the discussion. Situation A and Situation B are given on p. 16 of [Pereboom, 2001] as follows: “Situation A: Ms. Scarlet deliberately decides to kill Colonel Mustard at t_1 , and there are no factors beyond her control that deterministically produce her choice. When she chooses to kill the Colonel, she could have chosen not to kill him. There are no causal factors that would prevent her from not making the choice to kill Colonel Mustard. Situation B: Ms. Scarlet's choice to kill Colonel Mustard has precisely the same actual causal history as in A. But before she even started to think about killing Colonel Mustard, a neurophysiologist has blocked all the neural pathways not used in Situation A, so that no neural pathway other than the one employed in that situation could be used. Let us suppose that it is causally determined that she remain a living agent, and if she remains a living agent, some neural pathway has to be used. Thus every alternative for Ms. Scarlet is blocked except the one that realizes her choice to kill the Colonel. But the blockage does not affect the actual causal history of Ms. Scarlet's choice, because the blocked pathways would have remained dormant.”

on the event.”²¹ Consider, though, the following two-situation case discussed by Pereboom²², which seems to show that two situations with the (seemingly) same actual causal history may differ in that the one with (something relevantly similar to) a blockage may be essentially determined to proceed as it does:

Imagine a universe correctly described by Epicurean physics: At the most fundamental level all that exists is atoms and the frictionless void, and there is a determinate downward direction in which all atoms naturally fall – except if they undergo uncaused swerves.

Situation C: A spherical atom is falling downward through space, with a certain velocity and acceleration. Its actual causal history is indeterministic because at any time the atom can be subject to an uncaused swerve. Suppose that the atom can swerve in any direction other than upwards. In actual fact, from t_1 to t_2 it does not swerve.

A counterfactual situation diverges from C only by virtue of a device that eliminates alternative possibilities and all differences thereby entailed:

Situation D: The case is identical to C, except that the atom is falling downward through a straight and vertically oriented tube whose interior surface is made of frictionless material, and whose interior is precisely wide enough to accommodate the atom. The atom would not have swerved during this time interval, and the trajectory, velocity, and acceleration of the atom from t_1 to t_2 are precisely what they are in C.²³

So, what is the lesson from Situations C and D? One might initially want to claim that the causal histories in these two situations are identical between t_1 and t_2 , because the same thing happens to the particle during this time interval in both situations. This intuition, though, is questionable because it might be that the presence of the tube in Situation D makes the causal history of the particle in Situation D essentially deterministic. If the tube prevents the particle from swerving, and the possibility of swerving is what

²¹ Pereboom, 2001, p. 17.

²² Pereboom notes on p. 17 of Pereboom, 2001 that this two-situation case is “modeled on a reflection of Hunt’s” from Hunt’s personal correspondence with Fischer, cited on pp. 119-120 of Fischer, 1999.

²³ Pereboom, 2001, p. 17.

constitutes the indeterminate nature of the particle's movement, the tube precludes the movement of the particle from being indeterminate in nature; if the tube precludes the movement of the particle from being indeterministic, then the tube makes it the case that the particle's movement is essentially determined and, so, makes the particle's causal history deterministic in Situation D. In order for it to be the case that the movement of the particle through the tube between times t_1 and t_2 is indeterministic, it must be the case that at t_1 , t_2 , and any moments between (if there are any) the world could have been just as it was up until that moment, and yet the particle could have swerved (rather than not swerved, which is what it actually did). Given the placement of the tube, though, the particle could not have swerved and, so, its movement between t_1 and t_2 is not indeterministic. This would show that an object in two situations, here C and D, can (seemingly) have the same thing happen to it in both situations, and yet in one situation, C, what happens to the object is indeterministic, whereas in the other, D, the presence of a blocking mechanism (here, a tube) makes it the case that what happens to the object is determined. If this is the correct thing to say about situations C and D, this lesson applies to Hunt's blockage case in the following way: the intuition that the agent is morally responsible in Hunt's blockage case is taken to be reasonable, given that the actual causal history of the agent's choice is not affected by the presence of the blockages. The lesson from situations C and D is that the presence of such a blockage may affect the causal history (making it deterministic), even though the agent's choice would have been the same had the blockage not been present. Concerning what the lesson from situations C and D shows us about blockage cases, Pereboom says the following:

Sympathy for Frankfurt-style arguments is generated by the sense that moral responsibility is very much a function of the features of the actual causal history of

an action, to which restrictions that exist but would seem to play no actual causal role are irrelevant. However, in a scenario in which such restrictions, despite initial appearances, could be relevant to the nature of the actual causal history of an action after all, one's intuitions about whether the agent is morally responsible might become unstable. My own view is not that actual causal histories in blockage cases are clearly deterministic, but only that these considerations show that they may be. This type of problem should make one less confident when evaluating these difficult kinds of Frankfurt-style cases.²⁴

Internal-Sign Cases

Unlike No-Prior-Sign examples and Blockage examples, Internal-Sign examples do not attempt to eliminate a sign whose presence indicates that the agent will make the desired choice, *C*, and whose absence is sufficient for the intervention of the device to produce choice *C*. Rather, Internal-Sign examples are set up so that there is a sign of this sort, but, rather than occurring prior to the agent's action (i.e., rather than being a prior sign), the sign is internal to the event of the agent's choice/action. Such a case is given by Eleonore Stump in "Alternative Possibilities and Moral Responsibility: The Flicker of Freedom":

Suppose that a neurosurgeon Grey wants his patient Jones to vote for Republicans in the upcoming election. Grey has a neuroscope which lets him both observe and bring about neural firings which correlate with acts of will on Jones's part. Through his neuroscope, Grey ascertains that every time Jones wills to vote for Republican candidates, that act of will correlates with the completion of a sequence of neural firings in Jones's brain that always includes, near its beginning, the firing of neurons *a*, *b*, *c* (call this neural sequence "*R*"). On the other hand, Jones's willing to vote for Democratic candidates is correlated with the completion of a different neural sequence that always includes, near its beginning, the firings of neurons *x*, *y*, *z*, none of which is the same as those in neural sequence *R* (call this neural sequence "*D*"). For simplicity's sake, suppose that neither neural sequence *R* nor neural sequence *D* is also correlated with any further set of mental acts. Again, for simplicity's sake, suppose that Jones's only relevant options are an act of will to vote for Republicans or an act of will to vote for Democrats.

Then Grey can tune his neuroscope accordingly. Whenever the neuroscope detects the firing of *x*, *y*, and *z*, the initial neurons of neural sequence *D*, the neuroscope

²⁴ Pereboom, 2001, p.18.

immediately disrupts the neural sequence, so that it isn't brought to completion. The neuroscope then activates the coercive neurological mechanism which fires the neurons of neural sequence *R*, thereby bringing it about that Jones wills to vote for Republicans. But if the neuroscope detects the firing of *a*, *b*, and *c*, the initial neurons in neural sequence *R*, which is correlated with the act of will to vote for Republicans, then the neuroscope does not interrupt that neural sequence. It doesn't activate the coercive neurological mechanism, and neural sequence *R* continues, culminating in Jones's willing to vote for Republicans, without Jones's being caused to will this way by Grey.

And suppose that [in this case] Grey does not act to bring about neural sequence *R*, but that Jones wills to vote for Republicans, without Grey's coercing him to do so.²⁵

In setting up this case, Stump makes two presuppositions. The first presupposition is that there is a correlation between the mind and the brain, that mental events are somehow correlated with neural events. The nature of this correlation is left vague so that various different theories of mind/brain interaction can be accommodated by this Frankfurt-style case. This presupposition leaves it open whether each mental event is a temporally extended event that occurs over the time period from the beginning to the end of the corresponding neural event (in which case, the mental event does not occur unless the entire neural sequence occurs), or whether each mental event occurs after the completion of the corresponding neural event (in which case the mental event does not occur until the entire neural sequence occurs, when the last neuron firing of the neural event is completed). The second presupposition is that the correlation between a mental event and the neural event it is correlated with is a one-many relation; each mental event (say, a choice) is correlated to a neural event that consists of many neural firings.²⁶

²⁵ Stump, 1999a, pp. 303-305.

²⁶ "When I suddenly recognize my daughter's face across a crowded room, that one mental act of recognition, which feels sudden, or even instantaneous, is correlated with many neural firings as information from the retina is sent through the optic nerve, relayed through the lateral geniculate nucleus of the thalamus, processed in various parts of the occipital cortex, which take account of figure, motion, orientation in space, and color, and then processed further in cortical association areas. Only when the whole sequence of neural firings is complete, do I have the mental act of recognizing my daughter.

It seems to me that Stump's Frankfurt-style case is in no better shape than Mele and Robb's, and is faced with a similar problem (refer back to footnote 20). In Stump's case, it must be the case that there is indeterminacy in the event of the agent's choice being made such that this indeterminacy is the right kind of indeterminacy, meaning that (at least partly) in virtue of his choice being undetermined in this way the agent can be morally responsible for his choice. We can see that it must be the case that the indeterminacy in Stump's case lies in the initiation of the neural sequence that correlates with the mental event 'choosing to vote Republican' (this neural sequence begins with firings *a, b, c*), or in the initiation of the neural sequence that correlates with the mental event 'choosing to vote Democrat' (this neural sequence begins with firings *x, y, z*).²⁷ Stump wants to claim (see footnote 30) that unless a neural event is brought to completion, nothing mentally significant has occurred (i.e., if a neural event is terminated prior to its completion, no mental event occurs at all); so, for example, the firings of *a, b, c* alone are not mentally significant, because these firings are only the beginning of a neural event. If the indeterminacy occurs in the initiation of the neural sequence, and nothing mentally significant occurs until after the completion of the neural sequence, then

Whenever neural firings are associated with an act of will, I take it that in this case, as in all others, the correlation between the mental act and the firing of the relevant neurons is a one-many relation" [Stump, 1999a, p. 306].

²⁷ Stump intends for the indeterminacy to be located in the initiation of the neural sequence, as is noted in a later essay of hers: "So, as I originally presented the case, the initiating cause of the neural sequence in [the case] – whatever exactly that is – is itself indeterministic and to be understood in a way that doesn't preclude the indeterministic firing of the initial neurons of the sequence" [Stump, 2003, p. 142]. It must be that the indeterminacy in Stump's case is at no point later than the initiation of the neural sequence. It cannot be the case, for example, that the indeterminacy occurs after neurons *a, b, c* have fired, because then it would be the case that neurons *a, b, c* fire (in which case the device does not intervene) and yet a different neural event occurs, one other than that which correlates with the mental event 'choosing to vote Republican' (in which case the agent could choose otherwise, even though the intervening device is in place).

the indeterminacy is located prior to the occurrence of anything mentally significant.²⁸

For an agent's choice to be undetermined in the right kind of way, to allow for the agent to be morally responsible for that choice, it must be the case that up until the moment that choice is made, the past (which in Stump's case would include every neuron firing up until the last one of the neural sequence, everything up until something mentally significant occurs) could have been just as it was and yet it be possible for the agent to choose otherwise. This is not so in Stump's case, because once firings *a, b, c* occurs, the rest of the sequence will be completed and the mental event 'choosing to vote Republican' will occur. So, in Stump's Frankfurt-style case the agent's choice is not undetermined in the right way to allow for the agent to be morally responsible for this choice and, so, the conditions of an FR scenario are not met and, so, Stump's case is not a counterexample to PAP.

Necessary-Condition Cases

Necessary-Condition Cases make use of a device that relies on a prior sign in order to guarantee that the agent cannot choose otherwise than he does, but this prior sign is merely a necessary condition for the agent doing otherwise, rather than a sufficient condition of the agent doing otherwise. In a Necessary-Condition example, the agent (indeterministically) makes choice *C*, but in order for him to fail to make choice *C* he

²⁸ Suppose that Stump were to claim that the beginning of a neural sequence, for example firings *a, b, c*, is mentally significant and, so, the indeterminacy lies in the initiation of something mentally significant. This claim, though, would be problematic as well. If the beginning of a neural sequence is mentally significant, then the agent in Stump's case would have an alternative possibility: firings *x, y, z* could have happened, which would be the mentally significant beginning of the mental event 'choosing to vote Democrat' (perhaps the mental event in this case would be 'trying to vote Democrat'), in which case the device would intervene and bring about firings *a, b, c* and the mental event 'choosing to vote Republican' would also occur. Since there is this alternative possibility under the stipulation that the neural firings at the beginning of a neural sequence are mentally significant, under this stipulation Stump's case would not satisfy the conditions of an FR scenario and, so, would not be a counterexample to PAP.

would have has to (indeterministically) choose/do A. If the agent chooses/does A, though, the device would have intervened and brought about choice C. A Necessary-Condition example is given by Derk Pereboom, in its most revised version in “Defending Hard Incompatibilism.”

Tax Evasion (2): Joe is considering whether to claim a tax deduction for the substantial local registration fee that he paid when he bought a house. He knows that claiming the deduction is illegal, that he probably won’t be caught, and that if he is, he can convincingly plead ignorance. Suppose he has a very powerful but not always overriding desire to advance his self-interest regardless of the cost to others, and no matter whether advancing his self-interest involves illegal activity. Crucially, his psychology is such that the only way in this situation he could fail to choose to evade taxes is for moral reasons. His psychology is not, for example, such that he could fail to choose to evade taxes for no reason or simply on a whim. In addition, it is causally necessary for his failing to choose to evade taxes in this situation that he attain a certain level of attentiveness to these moral reasons. He can secure this level of attentiveness voluntarily. However, his attaining this level of attentiveness is not causally sufficient for his failing to choose to evade taxes. If he were to attain this level of attentiveness, Joe could, with his libertarian free will, either choose to evade taxes or refrain from so choosing (without the intervener’s device in place). More generally, Joe is a libertarian free agent. But to ensure that he chooses to evade taxes, a neuroscientist now implants a device, which, were it to sense the requisite level of attentiveness, would electronically stimulate his brain so that he would choose to evade taxes. In actual fact, he does not attain this level of attentiveness, and he chooses to evade taxes while the device remains idle.²⁹

In Tax Evasion 2, Joe (1) cannot choose otherwise than he does, choose to evade taxes, because if he met a necessary condition of choosing otherwise – attained a certain level of attentiveness to his moral reasons – the device would intervene and bring about the choice to evade taxes, (2) what brings about the situation described in (1), the device, in no way affects what Joe actually does, and (3) Joe is morally responsible for what he does because he did so on his own, was not determined to do so, etc.

Some proponents of PAP, though, may challenge (1) and claim that Joe does have a relevant alternative open to him: he can voluntarily choose to attain the requisite level of

²⁹ Pereboom, 2005, p. 231-232.

attentiveness to moral reasons for not evading taxes. His ability to so choose is a morally relevant alternative because had Joe chosen to reach the requisite level of attentiveness to moral reasons, the device would have intervened and forced Joe to choose to evade taxes. Under these circumstances, though, Joe would not be morally responsible for choosing to evade taxes, because this choice would be the product of the device and not made freely by Joe. Joe's ability to choose to attain this requisite level of attentiveness to moral reasons, then, is a relevant alternative – one that explains why Joe is morally responsible for his actual choice to evade taxes – because in this counterfactual scenario Joe is not morally responsible for choosing to evade taxes, while in the actual scenario he is morally responsible for choosing to evade taxes; thus, Joe has an alternative open to him in which he is not responsible for what in the actual scenario he chooses to do (i.e., he has an alternative open to him in which he is not blameworthy for what he does).³⁰

Pereboom, though, argues that having an alternative in which the agent is not blameworthy for what he does is not the right kind of alternative possibility that could explain why the agent is morally responsible for what he does; Pereboom argues that the alternative possibility that the objectors are referring to is not a robust alternate possibility and, so, his Tax Evasion 2 does satisfy the first condition of an FR scenario. According to Pereboom, in order for an alternative possibility to be robust it must satisfy the following conditions:

Robustness: For an alternative possibility to be relevant per se to explaining an agent's moral responsibility for an action it must satisfy the following characterization: she could have willed something other than what she actually

³⁰ Otsuka, 1998 and, Wyma, 1997.

willed such that she understood that by willing it she would thereby have been precluded from the moral responsibility she actually has for the action.³¹

So, according to Pereboom, in order for an alternative possibility to be the kind of possibility that is relevant to explaining an agent's moral responsibility for doing *R*, the alternative possibility must satisfy the two following conditions: (i) the agent could have chosen something, *Q*, such that *Q* is different from *R* AND the choosing of *Q* would thereby preclude the agent of the moral responsibility she has for what she actually chooses, *R*, and (ii) the agent understood that had she chosen *Q*, she would thereby be precluded from the moral responsibility she has for doing *R*. In Tax Evasion 2, the alternative possibility that is claimed by the objector to be morally relevant (the possibility of choosing to reach the requisite level of attentiveness, which would cause the device to intervene and produce the choice to evade taxes), while it satisfies condition (i) of Pereboom's definition of 'robustness,' does not satisfy condition (ii). Had Joe chosen to reach the requisite level of attentiveness to moral reasons, this choice would be different from his actual choices (Joe actually chooses to not pay attention to moral reasons and, so, chooses to evade taxes) and choosing to reach the requisite level of attentiveness would preclude him from the moral responsibility he has for his actual choices (as explained above, had Joe chosen to reach the requisite level of attentiveness, the device would have caused him to choose to evade taxes and Joe would not be morally responsible for this choice). This possibility of choosing to reach the requisite level of attentiveness, though, according to Pereboom, does not meet condition (ii) above:

But, Joe does not have an alternative possibility available to him that is robust. First, he does not even believe that if he had achieved the requisite level of attentiveness he would thereby have been precluded from responsibility for

³¹ Pereboom, 2005, p. 230.

deciding to evade taxes. For he believes that achieving this level of attentiveness is compatible with his never refraining from making this decision, or being seriously inclined so to refrain, and deciding to evade taxes instead. In addition, Joe does not know enough to understand that voluntarily achieving the requisite attentiveness would preclude him from responsibility for choosing to evade taxes. True, were he voluntarily to achieve this attentiveness, the intervention would take place, and he would not then have been responsible for this choice. Nevertheless, Joe does not understand that the intervention would then take place, or that as a consequence of this intervention he would be precluded from responsibility for choosing to evade taxes. Hence, no robust alternative possibility is available to him. Still, Joe is morally responsible for deciding to evade taxes.³²

Pereboom, then, maintains that Tax Evasion 2 meets the first condition of an FR scenario and, so, (since it meets the second and third conditions as well) is a counterexample to PAP. But, we can see that Pereboom's argument for the claim that Joe has no robust possibilities is dependent upon the correctness of his definition of 'robustness.' Pereboom's argument for the claim that Joe has no robust alternative possibilities in Tax Evasion 2 is just the following: the alternative possibility open to Joe does not meet my (Pereboom's) conditions on robustness, and, so, Joe has no robust alternative possibilities. Specifically, the alternative possibility open to Joe in Tax Evasion 2 does not meet what I will call Pereboom's epistemic condition on robustness, condition (ii) of the definition of 'robustness': in order for an alternative possibility to be robust the agent must have understood that by making a choice other than the one he actually makes he would thereby be precluded from the moral responsibility he has for what he actually does.

In the following section, I will argue that Pereboom's epistemic condition on robustness is not correct. While I will argue that there is some kind of an epistemic condition that must be included in a definition of 'robustness,' I will argue that

³² Pereboom, 2005, pp. 232-233.

Pereboom's epistemic condition demands too much of the agent in order for the alternative possibility open to him to be relevant to explaining the moral responsibility he has for what he actually does. If the definition of 'robustness' that I will propose is correct, the alternative open to Joe in Tax Evasion 2 is a robust alternative possibility. If, then, the definition of 'robustness' that I will propose is correct, Pereboom's Tax Evasion 2 will not meet the first condition of an FR scenario and, so, will not be a counterexample to PAP.

THE DEFINITION OF 'ROBUSTNESS'

In this section, I will argue that Pereboom's definition of 'robustness' is not correct, and given what I take is a more accurate account of 'robustness,' Pereboom's Frankfurt-style case, Tax Evasion 2, will leave the agent with robust alternative possibilities and, so, will not be a counterexample to PAP. If Pereboom's Tax Evasion 2 is not a counterexample to PAP, then the proponent of Frankfurt-style cases will either need to formulate a Necessary-Condition Frankfurt-style case that does meet all three conditions of an FR scenario, or will need to formulate a No-Prior-Sign, a Blockage, or an Internal-Sign example that meets the challenges previously made to such cases. To show that Pereboom's definition of 'robustness,' specifically his epistemic condition, is not correct, I will first look at his argument to support his epistemic condition, this argument being based upon our intuitions about a particular case. I will argue that our intuitions about this case do not need to be explained by positing Pereboom's epistemic condition on 'robustness,' but, rather, are better explained by the epistemic condition(s) that I will suggest. I will spend the remainder of this essay discussing the condition(s) that I propose, and how given the definition of 'robustness' that I propose, Pereboom's Tax Evasion 2 is not a counterexample to PAP. Before taking a look at Pereboom's argument for his epistemic condition on 'robustness,' I will briefly discuss the criteria of adequacy for (or, what we are looking for in) a definition of 'robustness.'

The term 'robustness' is a technical term that picks out a certain feature (or, a set of features) of some, but not all, of the counterfactual/alternative possibilities that we

believe are open to agents. We will say that an agent has a robust alternative possibility if and only if there is a counterfactual sequence open to him that has this certain feature (or set of features). In order for a definition of ‘robustness’ to be an adequate definition, it must spell out the necessary and sufficient conditions for when some agent has a robust alternative possibility, and the definition must do this in a way that accords with our intuitions about particular cases. Intuitively, it seems to me that an agent has a robust alternative possibility in those cases in which the addition or removal of this possibility affects our judgment of the agent’s moral responsibility. It seems to me that our intuitions concerning when some alternative possibility is robust are guided, then, in the following way: An agent has a robust alternative possibility to choosing/doing x in those cases in which the removal of that alternative possibility would change our judgment of him from morally responsible for choosing/doing x to not morally responsible for choosing/doing x ; and, an alternative possibility to x is robust in those cases in which the addition of that alternative possibility would change our judgment of an agent from not morally responsible for choosing/doing x to morally responsible for choosing/doing x .

It is the job of one who proposes a definition of ‘robustness’ to give the necessary and sufficient conditions that must be met in order for an agent to have a robust alternative possibility, and the satisfaction of these conditions should occur when and only when our intuition, guided in the way outlined above, tells us that the agent has an alternative open to him that affects our judgment of his moral responsibility. In what follows, I will argue that Pereboom’s argument for his definition of ‘robustness’ is flawed, and I will propose a definition of ‘robustness’ that will do a better job of explaining our intuitions about an agent’s moral responsibility in particular cases.

In arguing for his epistemic condition on robustness, Pereboom gives a case in which (a) (the agent presumably knows that) there was an alternative open to the agent, that he could have chosen otherwise, and (b) it is clear that the agent is morally responsible, and (c) the agent does not understand that if he had chosen otherwise he would thereby be precluded from the moral responsibility he has for what he does. Roughly, Pereboom's argument for his epistemic condition on robustness is the following: in this case, (a) the agent's ability to choose otherwise is not relevant to (b) the agent's moral responsibility for what he does because (c) the agent does not understand that by choosing otherwise he would thereby be precluded from the responsibility he has for what he does:

Robustness also has an epistemic dimension, and it is important that it be made explicit in the characterization of this notion. Imagine that the only way in which Jones could have voluntarily avoided deciding to kill Smith is by taking a sip from his coffee cup prior to making the decision, and this is only because it was poisoned so that taking a sip would have killed him instantly. Suppose that Jones does not understand that this action would preclude his deciding to kill, because he has no idea that the coffee is poisoned. In this situation, Jones could have voluntarily behaved in such a manner that would have precluded the action for which he was in fact blameworthy, as a result of which he would have avoided the moral responsibility he actually has. But whether he could have voluntarily taken the sip from the coffee cup, not understanding that it would render him blameless in this way, is irrelevant *qua* alternative possibility to explaining why he is morally responsible for deciding to kill. Despite the fact that Jones could have voluntarily taken a sip from his coffee cup, and doing so would have rendered him not morally responsible for deciding to kill, this alternative possibility is nevertheless insufficiently robust to have an important role in grounding the agent's moral responsibility.¹

So, Jones (b) is morally responsible for choosing to kill Smith, despite the fact that (a) Jones could have voluntarily taken a sip of poisoned coffee, which would thereby preclude him from the moral responsibility he has for killing Smith (had Jones

¹ Pereboom, 2005, p. 230.

voluntarily chosen to take a sip of poisoned coffee, he would die and, so, he would not kill Smith and, so, he would not be morally responsible for killing Smith). The reason that (a) Jones's having the ability to choose otherwise – to choose to take a sip of coffee – is irrelevant to explaining (b) Jones's moral responsibility for what he does is because (c) Jones did not understand that had he chosen otherwise he would be precluded from the moral responsibility he has for killing Smith. Jones does not know that his coffee is poisoned and, so, does not understand that by taking a sip he would thereby be precluded from the moral responsibility he has for his choice to kill Smith. What this suggests, according to Pereboom, is that in order for an alternative possibility to be relevant to explaining an agent's moral responsibility, it must be the case that the agent understands that by choosing otherwise he would thereby be precluded from the responsibility he has for what he does; in order for an alternative possibility to be robust, it must be the case, according to Pereboom, that the agent understands that he could choose otherwise and understands that by choosing otherwise he would thereby be precluded from the moral responsibility he has for what he actually does.

It seems to me that Pereboom's reasoning described above is flawed in that (c), that the agent does not understand that by choosing otherwise he would thereby be precluded from the moral responsibility he has for what he does, is not the only feature of the case that could be appealed to in order to explain why (a) the agent's ability to choose otherwise is irrelevant to (b) the agent's moral responsibility. There is another feature of this case given above, a feature that better explains why (a) the agent's ability to choose otherwise is irrelevant to (b) the agent's moral responsibility: (d) the agent does not have an alternative (to what we can assume he is being claimed morally responsible for,

choosing to kill and killing Smith) open to him such that he believes that as a result of choosing otherwise, the likelihood will be different (than if he did not choose otherwise) that he choose to kill and kill Smith. I believe that feature (d) of the case more plausibly explains why (a) Jones's ability to choose otherwise is irrelevant to (b) his moral responsibility for choosing to kill and killing Smith. It is true that by choosing to take a sip of coffee, Jones would not choose to kill Smith (and would not kill Smith), but Jones does not believe that this is the case, nor should we say he has any good reason to believe so. This explains why Jones's ability to choose otherwise is irrelevant to his moral responsibility for what he chooses and does. Given this analysis of Pereboom's poisoned coffee case, I propose the following definition of 'robustness':

Robustness: an agent has a robust possibility, y , to x at t_n (where t_n is not necessarily the time x was chosen/done) if and only if (i) had he done y at t_n he would not have done x and, (j) he believes that he could have chosen and done otherwise, y , at t_n and, (g) the agent believes that he can choose/do y at t_n such that the likelihood that he choose/do x is different than it would be if he does not choose/do y at t_n .²

So, the reason that Jones's ability to choose to take a sip of coffee is irrelevant to his moral responsibility for choosing to kill and killing Smith is because Jones does not believe that taking a sip will change the likelihood that he chooses to kill and kills Smith. According to the proponent of PAP, then, I am suggesting that unless Jones has an alternative to killing Smith that meets the criteria given above, Jones is not morally responsible for killing Smith in the poisoned coffee case.

² Recall that the reformulated version of PAP is as follows: having a robust alternative possibility to choosing/doing x is necessary for an agent to be morally responsible for choosing/doing x . With this definition of 'robustness,' then, what I am suggesting is that the intuition behind PAP really amounts to the following: an agent is morally responsible for choosing/doing x at t_1 only if he could have either (a) chosen/done otherwise at t_1 or, (b) chosen/done otherwise, y , at some time prior to t_1 , where y satisfies my definition of 'robustness.'

Let's test this definition against our intuition, guided in the way previously outlined, about Jones's moral responsibility in a slightly modified version of the poisoned coffee case. In Poisoned Coffee Case 2, Jones chooses to not and does not take a sip of coffee and then chooses to kill and does kill Smith. Unbeknownst to Jones, though, his coffee is poisoned and, so, had he taken a sip he would have died instantly. Suppose, though, that Jones could not have done anything other than choose to not and not take a sip of coffee and then choose to kill and kill Smith. Suppose that, unbeknownst to Jones, he has a mental illness such that he always acts on the first course of action that comes to his mind (perhaps he has some kind of attention deficit disorder that prevents him from going through an adequate deliberative process of weighing reasons for and against various courses of action), and in this case the first and only course of action he considers is to not take a sip of coffee and then go kill Smith. Suppose also that, unbeknownst to Jones, there is a device in his brain that is indeterministically stimulating his brain in such a way that the device is indeterministically supplying his brain with the first course of action that he considers, and, so, as a result of the combination of this device and Jones's mental illness, Jones cannot do anything other than carry out the course of action supplied to him by the device. In Poisoned Coffee Case 2, the device supplied Jones with the choice to not take a sip of coffee and the choice to kill Smith and, so, Jones follows through with this course of action and does not take a sip of coffee and does kill Smith. It seems plausible that in Poisoned Coffee Case 2 Jones is not responsible for choosing to kill Smith.

Now consider Poisoned Coffee Case 3, which is just like Poisoned Coffee Case 2, except that the device is not present in Poisoned Coffee Case 3. In Poisoned Coffee Case

3 Jones could have taken a sip of poisoned coffee had he chosen to do so, i.e., had this been the first course of action he considered (in which case, he would not choose to kill and would not kill Smith, because he would have died instantly). In Poisoned Coffee Case 3, it is also the case that Jones believes that taking a sip of coffee will make it less likely that he chooses to kill and kills Smith (i.e., suppose Jones has an alternative possibility that satisfies my proposed definition of ‘robustness’); perhaps Jones believes this based on past experiences with drinking coffee and making important decisions (such as the decision to kill someone) and past experience with drinking coffee and shooting – Jones has previously taken a sip of coffee before practicing his shooting and this has in the past reduced his accuracy. (It is still the case, though, that Jones, due to his mental illness, can only act on the first course of action he considers, where in this case the first course of action he considers is not taking a sip of coffee and then killing Smith). In Poisoned Coffee Case 3, the case in which Jones has an alternative possibility that satisfies my proposed definition of ‘robustness,’ I believe that the intuition will be that Jones is responsible for choosing to kill Smith whereas he is not responsible for doing this in Poisoned Coffee Case 2, the case in which he had no such alternative. That Jones has an alternative possibility such that he believes that by choosing otherwise he will be less likely to do something – something that he knows is wrong, choosing to kill Smith – affects our judgment about whether or not Jones is morally responsible for what he does. If I am right that one will have the intuition that Jones is morally responsible for choosing to kill and killing Smith in Poisoned Coffee Case 3, and I am right that one will have the intuition that he is not morally responsible for this in Poisoned Coffee Case 2, and the difference between these two cases is that in Poisoned Coffee Case 2 Jones does not and

in Poisoned Coffee Case 3 Jones does have an alternative open to him that satisfies my proposed definition of 'robustness,' then the addition of such an alternative to a case affects our judgment of an agent's moral responsibility in the outlined above (i.e., the addition of an alternative that satisfies my definition of 'robustness' changes our judgment from not morally responsible for choosing and doing x , to morally responsible for choosing and doing x).

One might object that in Poisoned Coffee Case 3, it is still the case that Jones has a mental illness that makes it the case that he can only consider one course of action at a time. One might claim that, despite having a robust alternative possibility, since Jones has this mental illness, it is unjustifiable to hold him morally responsible for what he does. Since Jones is not morally responsible in the case in which he is given an alternative possibility that satisfies my definition of 'robustness,' my definition of 'robustness' is not correct.

I believe that this objection is mistaken, and that there is no reason that an agent with a mental illness cannot be held morally responsible for choices and actions when that agent has an alternate possibility to what he does and the agent meets a certain epistemic condition concerning the alternative(s) open to him. If an agent could have chosen otherwise and believes that he could do (could have done) so such that choosing otherwise will (would have) make (made) it less likely that he kill another person, it is reasonable to hold him responsible for killing that person. In Poisoned Coffee Case 3, even though Jones has a mental illness, this mental illness does not prevent Jones from choosing otherwise; in Poisoned Coffee Case 3, Jones can choose otherwise than to not take a sip of coffee and to kill Smith. It could have been the case that the first course of

action Jones considered was to take a sip of coffee and then not kill Smith, in which case Jones would have taken a sip of poisoned coffee and then died (and, so, not killed Smith). It seems, then, that an agent that has a mental illness will only be exempt from moral responsibility for doing something *because of that mental illness*, if it is the case that his mental illness either prevents him from doing anything other than what he does, or if it prevents the agent from having the belief associated with condition (g) of my definition of ‘robustness.’ The mental illness Jones has in Poisoned Coffee Case 3 neither prevents him from doing anything other than what he does, nor does it prevent him from having the belief associated with condition (g) of my definition of ‘robustness’ and, so, Jones’s mental illness will not exempt him from moral responsibility for choosing to kill and killing Smith.

So, in Poisoned Coffee Case 3, Jones is morally responsible for choosing to not take a sip of coffee, choosing to kill Smith, and killing Smith, because he has a robust alternative possibility to choosing and doing these things: he could have chosen to take a sip of coffee such that (j) if he chose to take a sip of coffee he would not have chosen to not take a sip of coffee, chosen to kill Smith, or killed Smith, and (g) he believed that choosing to take a sip of coffee would make it less likely that he choose to kill Smith and kill Smith. Notice, though, that according to Pereboom’s definition of ‘robustness’ Jones would not have a robust alternate possibility to choosing to kill and killing Smith in Poisoned Coffee Case 3. Recall that according to Pereboom, in order to have a robust alternative possibility to what an agent does, that agent must be able to choose otherwise such that he understands that by choosing otherwise he would be precluded from the responsibility he has for what he actually does. So, in Poisoned Coffee Case 3, in order

for Jones's ability to take a sip of coffee to be a robust alternate possibility to choosing to kill and killing Smith, according to Pereboom, Jones would need to understand that by taking a sip of coffee he would be precluded from the moral responsibility he has for choosing to kill and killing Smith. Jones does not, though, understand this in Poisoned Coffee Case 3, because Jones believes only that the likelihood of choosing to kill and killing Smith is reduced if he takes a sip of coffee. In Poisoned Coffee Case 3 Jones believes that if he takes a sip of coffee, it is still possible he will choose to kill and kill Smith. So, since Jones does not understand that by taking a sip of coffee he will be precluded from the moral responsibility he has for choosing to kill and killing Smith, Jones does not, according to Pereboom's definition of 'robustness,' have a robust alternative possibility. If, then, one agrees that Jones has an alternate possibility that is relevant to his moral responsibility for choosing to kill and for killing Smith in Poisoned Coffee Case 3, one cannot accept Pereboom's definition of 'robustness.' I believe that one will have the intuition that Jones has an alternative open to him in Poisoned Coffee Case 3 that is relevant to his moral responsibility for choosing to kill and for killing Smith and, so, one will see that my proposed definition of 'robustness' is more accurate than Pereboom's definition of 'robustness.'

Given the arguments presented above for the definition of 'robustness' I am proposing, we can see how Pereboom's Tax Evasion 2 Frankfurt-style case leaves the agent with a robust alternative possibility and, so, is not a counterexample to PAP because it does not satisfy the first condition on an FR scenario. As given, though, Pereboom's Tax Evasion 2 is underdescribed and, so, I will make a stipulation about the agent's psychology and show that, given this stipulation, the agent has a robust

alternative possibility. Recall that in Tax Evasion 2, Joe chooses to not pay attention to his moral reasons for not evading taxes and, so, chooses to and does evade his taxes. The only way that Joe could have not evaded his taxes was by choosing to pay sufficient attention to his moral reasons for not evading. Achieving this requisite level of attentiveness to his moral reasons is necessary but not sufficient for his choosing to not evade taxes; had he reached this requisite level of attentiveness (without the device being in place) he could either choose to evade or fail to choose to evade. It is the case, though, that a device is in place such that if Joe reaches the requisite level of attentiveness to his moral reasons for not evading, the device will intervene and bring about the choice to evade and Joe would evade his taxes. Joe actually chooses to not pay attention to his moral reasons and, so, chooses to evade and does evade, even though Joe could have chosen to pay attention to his moral reasons for not evading taxes. Recall that at this point Pereboom argues that this alternative open to Joe – the ability to choose to pay sufficient attention to moral reasons - is not a robust alternative to what he actually does. Given the following reasonable supposition I will make about Joe’s psychology, if the definition of ‘robustness’ I have proposed is correct, then Joe’s available option of choosing to pay sufficient attention to his moral reasons for not evading taxes is a robust alternative possibility to his choosing to not pay attention to moral reasons: Joe believes that he could have chosen otherwise - chosen to pay sufficient attention to his moral reasons – such that if he chose to pay sufficient attention to his moral reasons, he would have been less likely to evade his taxes. I will call this version of Tax Evasion 2, in which Jones has a robust alternate possibility to choosing to not pay attention to moral reasons, Tax Evasion 2^w. In Tax Evasion 2^w Joe’s ability to choose to pay sufficient attention to his

moral reasons is an alternative to what he actually chooses, the availability of which is relevant to explaining his moral responsibility for choosing to not pay attention to his moral reasons.

So, in Tax Evasion 2, the defender of PAP can argue that Joe is morally responsible for choosing to not pay attention to his moral reasons (on the supposition that Tax Evasion 2 is just like Tax Evasion 2^ψ), because he had a robust alternate possibility to doing so. Perhaps, though, this is all Joe is morally responsible for in Tax Evasion 2; perhaps Joe is not morally responsible for choosing to evade and for evading his taxes. After reading Tax Evasion 2, I believe one has the intuition that Joe is responsible for something, and the most obvious candidate for this is the action he performs – evading his taxes. It seems to me, though, that it is open to the proponent of PAP to claim that this is not what Joe is morally responsible for in Tax Evasion 2, and that he is only morally responsible for his choice to not pay attention to moral reasons, in which case Tax Evasion 2 is not a counterexample to PAP because Joe does have a robust alternative possibility to what he is morally responsible for, choosing to not pay attention to his moral reasons.

What, though, would be the case if Joe does not have such a belief, if Joe does not believe that choosing to pay sufficient attention to his moral reasons will make it less likely that he choose to evade taxes? It seems very hard to imagine that a person could not have any beliefs concerning how the consideration of moral reasons will affect one's choices and behavior. Consider case Tax Evasion 2^{~ψ}, in which we suppose that Joe does not have any such beliefs - that he does not believe that paying sufficient attention to his moral reason will in any way change the likelihood that he will evade taxes – would we

still want to claim that Joe is morally responsible for anything he does? If we would want to claim that Joe is morally responsible for choosing to not pay attention to his moral reasons and for choosing to evade and evading his taxes, despite the fact that he does not, according the definition of ‘robustness’ I have proposed, have a robust alternate possibility to choosing/doing as such, then either my definition of ‘robustness’ or PAP (or both) would be false. I think, though, that both my definition of ‘robustness’ and PAP can be upheld in the case at hand, the case in which Joe’s alternative possibility – the ability to choose to pay sufficient attention to his moral reasons – does not satisfy my definition of ‘robustness.’

Tax Evasion 2^{~ψ} is just like the original poisoned coffee case given by Pereboom, even though it is very difficult for us to imagine that any person could hold the choice of whether or not to consider moral reasons as on par (with respect to the effect that such consideration will have on one’s future choices and actions) with the consideration of whether or not to take a sip of coffee. If we can picture that, with respect to the effect that the agent believes that the consideration of whether or not to x will have on his future choices and actions, Joe believes that the choice of whether or not to pay sufficient attention to his moral reasons is on par with the choice of whether or not to take a sip of coffee (this is what Joe believes in Tax Evasion 2^{~ψ}), I believe the intuition that Joe is morally responsible for choosing to not pay sufficient attention to his moral reasons, choosing to evade, and evading taxes will fade away. In what follows, I will argue that in Tax Evasion 2^{~ψ}, Joe is neither responsible for choosing to not pay sufficient attention to his moral reasons, nor for choosing to evade and evading his taxes.

Consider Tax Evasion 2* which is set up just like Tax Evasion 2, except that Joe* will choose to evade and will evade taxes only unless he chooses to and takes a sip of coffee; if Joe* does choose to and take a sip of coffee, under normal circumstance, he could then choose whether or not to evade taxes. Suppose, though, that there is a device such that if Joe* will choose to take a sip of coffee, the device will prevent him from making this choice and will bring about the choice to not take a sip. Joe*, then, cannot choose to do anything but not take a sip of coffee and evade his taxes. Suppose further that Joe* does not believe that the decision of whether or not to take a sip of coffee will in any way affect the likelihood of making any future choice or performing any future action. Suppose that for his own reasons, Joe* does not take a sip of coffee and chooses to evade and does evade his taxes. In Tax Evasion 2*, I cannot see what could justify the claim that he is morally responsible for not taking a sip of coffee. Taking or not taking a sip of coffee is not normally the kind of activity that one is *morally* responsible for, unless of course, the agent has some kind of a belief associated with the effects of taking or not taking a sip on his behavior (and, Joe* does not have any beliefs like this in Tax Evasion 2*). If Joe* is not morally responsible for not taking a sip of coffee in Tax Evasion 2*, and if the decision of whether or not to take a sip of coffee is considered by the agent as on par (with respect to the affect on one's future choices and behavior) with the decision of whether or not to consider moral reasons, then Joe is not morally responsible for choosing to not pay attention to his moral reasons in Tax Evasion 2^{~ψ}, the case in which he does not believe that paying attention to his moral reasons will affect the likelihood that he evade taxes. So, if Joe did not have a robust, according to my definition of 'robust,' alternative to choosing to not pay attention to his moral reasons (as he would

not in Tax Evasion 2^{~ψ}), he would not be morally responsible for choosing to not pay attention to his moral reasons.

A defender of Frankfurt-style cases might object at this point by claiming that the previous argument has suggested a Frankfurt-style case that satisfies all three conditions of an FR scenario: Tax Evasion 2^{~ψ}. In Tax Evasion 2^{~ψ}, Joe does not have a robust alternate possibility to choosing to evade and evading his taxes (given the placement of the device Joe cannot do anything but choose to evade and evade taxes), what makes it the case that he cannot do otherwise does not bring about his actual choice/action (in Tax Evasion 2^{~ψ}, it is still the case that Joe chooses to not pay attention to his moral reasons and, so, chooses to evade and evades his taxes, and he chooses/does so without the intervention of the device), and it is plausible to claim that Joe is morally responsible for choosing to evade and for evading taxes since he did so indeterministically (just like Joe did so indeterministically in Tax Evasion 2).

I do not, though, believe that Tax Evasion 2^{~ψ} is a case in which the indeterminacy in Joe's choice to evade taxes is the right kind of indeterminacy to allow for moral responsibility for so choosing and doing. We must remember what role Tax Evasion 2 (or any Necessary-Condition Frankfurt-style case) is playing in the debate. Recall that Pereboom's Tax Evasion 2, and any Necessary-Condition example for that matter (for example, Tax Evasion 2^{~ψ}), is a response on behalf of the proponent of Frankfurt-style cases to the second horn of the dilemma defense of PAP. The challenge set forth in the second horn of the dilemma is to give a case in which the agent's choice is *in some relevant way undetermined* and yet the agent has no robust alternative possibilities. The whole motivation for challenging the proponent of Frankfurt-style cases in this way

comes from the Incompatibilist who claims that in order for an agent to be morally responsible for what he chooses or does, his choice must be in some relevant way undetermined. The Incompatibilist who makes such a challenge, though, will not be satisfied with the presentation of an agent whose choice is undetermined in just any way.

For example, consider a case in which an agent's choice is undetermined in just the following way: a particular agent who is 21 years old will at age 24 choose to go to medical school only unless at age 22 he chooses that red is his favorite color. Suppose that the agent does choose at age 24 to go to medical school. The Incompatibilist, though, will not accept that the kind of indeterminacy present in this case is the right kind of indeterminacy to allow for moral responsibility, and I believe that the Incompatibilist will not accept this kind of indeterminacy for the following reason: the agent has no reason to believe that there is any connection between his choice at age 22 of a favorite color and his choice at age 24 of what career to pursue. In order for the indeterminacy present in a case to be the kind of indeterminacy required for an agent to be morally responsible for some choice *C*, it must be the case that whatever explains the indeterminacy (here, the ability to choose or not choose red as his favorite color) is in some way relevant to the undetermined choice for which the agent is claimed to be morally responsible (here, the choice to go to medical school). Contrast the kind of indeterminacy present in the case just given with the indeterminacy present in Pereboom's Tax Evasion 2 case. In Pereboom's case, what explains the fact that Joe's choice concerning whether to evade taxes is undetermined is that he has the option of (voluntarily) paying attention to moral reasons for and against doing so. The reason, I believe, that this kind of indeterminacy will satisfy the Incompatibilist proponent of the dilemma defense of PAP is because the

weighing of moral reasons for and against doing something is taken to be (assumed to be) relevant to the undetermined choice Joe is claimed to be morally responsible for, choosing to evade taxes. Although the kind of indeterminacy present in Tax Evasion 2 is the right kind *to allow for* the agent to be morally responsible for what he does, it is not necessarily the case that this kind of indeterminacy suffices for moral responsibility and, so, one can still argue that Joe is not morally responsible for choosing to evade and for evading taxes in Pereboom's Tax Evasion 2. (It is still the case, then, that it is open to the proponent of PAP to argue that Joe is not morally responsible for choosing to evade and for evading taxes in Tax Evasion 2, and that what he is morally responsible for is choosing to not pay sufficient attention to moral reasons, something he has a robust alternative to choosing/doing.).

The indeterminacy in Tax Evasion 2^{~ Ψ} , the case in which Joe has no belief concerning the effect, on his future choices and actions, of paying sufficient attention to his moral reasons is like the indeterminacy in the case of the agent who will choose at age 24 to go to medical school only unless he chooses at age 22 that red is his favorite color. In the case in which Joe has no belief concerning the effect, on his future choices and actions, of paying sufficient attention to his moral reasons, what explains that his choice to evade taxes is undetermined is that he could fail to so choose if he chooses to pay sufficient attention (if the device were not in place). But, if his belief concerning the effect, on his future choices and actions, of considering moral reasons is on par with his belief concerning the effect, on his future choices and actions, of considering whether or not to take a sip of coffee, then the indeterminacy present in Tax Evasion 2^{~ Ψ} is just like the indeterminacy present in the case of the student who will choose at age 24 to go to

medical school only unless he chooses at age 22 that red is his favorite color. If it is the case, then, that Joe does not believe that paying attention to moral reasons will affect the likelihood of choosing to evade and evading taxes, there will not be the right kind of indeterminacy present in the case to allow for moral responsibility. Tax Evasion 2^ψ (i.e., the case in which Joe does not have an alternative open to him that satisfies my proposed definition of ‘robustness’), then, is not a case that is a response to the dilemma defense of PAP and, so, will not help the proponent of Frankfurt-style cases argue his way out of the dilemma.

In Tax Evasion 2^ψ, then, Joe is neither morally responsible for choosing to not pay sufficient attention to moral reasons, nor is he morally responsible for choosing to evade and for evading taxes. That Joe has no robust alternative in Tax Evasion 2^ψ to so choosing/doing according to my proposed definition of ‘robustness,’ then, is neither problematic for my proposed definition of ‘robustness,’ nor for PAP.

On the one hand, then, we can see that if Joe has an alternative possibility open to him that satisfies my proposed definition of ‘robustness,’ (the version called Tax Evasion 2^ψ) Tax Evasion 2 will not satisfy the first condition of an FR scenario (the agent must have no robust alternative possibilities to what he actually chooses to do) and, so, will not be a counterexample to PAP. On the other hand, if the alternative possibility open to Joe does not meet my proposed definition of ‘robustness,’ (the version called Tax Evasion 2^ψ) Tax Evasion 2 will not satisfy the third condition of an FR scenario (the agent must be morally responsible for his actual choice) and, so, will not be a counterexample to PAP. I conclude, then, that Pereboom’s Frankfurt-style case does not satisfy all three conditions of an FR scenario and, so, is not a counterexample to PAP.

The discussion of section three leads me to conclude that the proponents of Frankfurt-style cases have yet to give a case that is a counterexample to PAP because they have yet to give a case that satisfies all three conditions of an FR scenario. It is open to the proponent of Frankfurt-style cases to support Pereboom's definition of 'robustness' or some other definition that shows that Tax Evasion 2 does not include such alternatives open to the agent. The proponent of Frankfurt-style cases could also support one of the three other styles of cases – the No-Prior-Sign example, the Blockage example, or the Internal-Sign example – by showing that something is mistaken with the criticisms of such cases presented in the second section of this essay. It is also open to the proponent of Frankfurt-style cases that finds the definition of 'robustness' I have here proposed attractive to give a Necessary-Condition example in which the agent's choice is undetermined in the right kind of way and yet the agent has no robust alternative possibilities given the definition of 'robustness' that I have proposed.

In this essay I have argued that the proponent of Frankfurt-style cases has yet to give a case that satisfies all three conditions of an FR scenario and, so, has yet to give a case that is a counterexample to PAP. I have argued this by discussing a dilemma that is posed to the proponent of Frankfurt-style cases, the dilemma defense of PAP. I have argued that the proponent of Frankfurt-style cases has not adequately responded to the first horn of the dilemma; if it is supposed that the prior sign is sufficient for the agent's choice in a Frankfurt-style case, the use of a Frankfurt-style case to support the Compatibilist's conclusion is unnecessary and, so, this first horn is problematic if the Compatibilist claims that Frankfurt-style cases go some way to advancing the debate, between Compatibilists and Incompatibilists, concerning moral responsibility. Although

the problem with the supposition that the prior sign suffices for the agent's choice has, up until now, been posed as a problem of begging the question against the Incompatibilist, I have argued that this is not the real problem with the supposition but that, regardless, the proponent of Frankfurt-style cases (Fischer) has not adequately responded to the real problem posed by the supposition of the first horn of the dilemma defense of PAP.

I have also argued that the proponent of Frankfurt-style cases has not adequately responded to the second horn of the dilemma defense of PAP. I have done this by discussing the four types of cases that proponents of Frankfurt-style cases have used to respond to this horn of the dilemma, showing that neither strategy succeeds in satisfying all three conditions of an FR scenario and, so, neither strategy succeeds in giving a counterexample to PAP. I have paid special attention to the Necessary-Condition strategy, used by Derk Pereboom, showing that the proponent of PAP can use my proposed definition of 'robustness' to argue that Pereboom's Tax Evasion 2 is not a counterexample to PAP. I have gone some way to showing that my proposed definition of 'robustness' is more accurate than Pereboom's definition of 'robustness,' and, since little attention has been paid in the literature to attempting to define this term, I have (hopefully) helped us to understand more about the Principle of Alternate Possibilities and when it is satisfied.

LIST OF REFERENCES

- John Martin Fischer, *The Metaphysics of Free Will: An Essay on Control* (Cambridge, MA.: Blackwell, 1994).
- John Martin Fischer, "Recent Work on Moral Responsibility," *Ethics* 110 (1999): 93-139.
- John Martin Fischer, "Frankfurt-type Examples and Semi-Compatibilism," in *The Oxford Handbook of Free Will*, ed. Robert Kane, ed. (New York: Oxford University Press, 2002), 281-308.
- John Martin Fischer, "Free Will and Moral Responsibility," in *The Oxford Handbook of Ethical Theory*, ed. David Copp (New York: Oxford University Press, 2006), 321-354.
- Harry G. Frankfurt, "Alternate Possibilities and Moral Responsibility," *Journal of Philosophy* 66 (1969): 829-839.
- Harry G. Frankfurt, "Freedom of the Will and the Concept of a Person," *Journal of Philosophy* 68 (1971): 5-20.
- Stewart Goetz, "Stumping for Widerker," *Faith and Philosophy* 16 (1999): 83-89.
- Stewart Goetz, "Frankfurt-style Counterexamples and Begging the Question," in *Midwest Studies in Philosophy Volume XXIX: Free Will and Moral Responsibility*, eds. Peter A. French and Howard K. Wettstein, guest ed. John Martin Fischer (Boston: Blackwell Publishing, 2005), 83-105.
- David Hunt, "Moral Responsibility and Unavoidable Action," *Philosophical Studies* 97 (2000): 195-227.
- Robert Kane, *Free Will and Values* (Albany: State University of New York Press, 1985).
- Robert Kane, *The Significance of Free Will* (New York: Oxford University Press, 1996).
- Robert Kane, "Responses to Bernard Berofsky, John Martin Fischer and Galen Strawson," *Philosophy and Phenomenological Research* 60 (2000): 157-167.
- Robert Kane, "Responsibility, Indeterminism and Frankfurt-style Cases: A Reply to Mele and Robb," in *Moral Responsibility and Alternate Possibilities*, eds. David Widerker and Michael McKenna (Burlington, VT: Ashgate, 2003), 91-105.

- Michael McKenna, "Alternative Possibilities and the Failure of the Counterexample Strategy," *Journal of Social Philosophy* 28 (1997): 71-85.
- Alfred R. Mele and David Robb, "Rescuing Frankfurt-Style Cases," *The Philosophical Review* 107 (1998): 97-112.
- Michael Otsuka, "Incompatibilism and the Avoidability of Blame," *Ethics* 108 (1998): 685-701.
- Derk Pereboom, *Living Without Free Will* (New York: Cambridge University Press, 2001).
- Derk Pereboom, "Defending Hard Incompatibilism" in *Midwest Studies in Philosophy Volume XXIX: Free Will and Moral Responsibility*, eds. Peter A. French and Howard K. Wettstein, guest ed. John Martin Fischer (Boston: Blackwell Publishing, 2005), 228-247.
- P.F. Strawson, "Freedom and Resentment," in *Free Will*, ed. Gary Watson (New York: Oxford University Press, 1982), 59-80.
- Eleonore Stump, "Alternative Possibilities and Moral Responsibility: The Flicker of Freedom," *The Journal of Ethics* 3 (1999a): 299-324.
- Eleonore Stump, "Dust, Determinism, and Frankfurt: A Reply to Goetz," *Faith and Philosophy* 16 (1999b): 413-422.
- Eleonore Stump, "Moral Responsibility without Alternative Possibilities," in *Moral Responsibility and Alternative Possibilities*, eds. David Widerker and Michael McKenna (Burlington, VT: Ashgate, 2003), 139-158.
- Peter van Inwagen, *An Essay on Free Will* (New York: Oxford University Press, 1983).
- David Widerker, "Libertarianism and Frankfurt's Attack on the Principle of Alternate Possibilities," *Philosophical Review* 104 (1995): 247-261.
- David Widerker, "Frankfurt's Attack on the Principle of Alternative Possibilities: A Further Look," *Philosophical Perspectives* 14 (2000): 181-201.
- David Widerker, "Responsibility and Frankfurt-type Examples," in *The Oxford Handbook of Free Will*, ed. Robert Kane (New York: Oxford University Press, 2002) 323-336.
- David Widerker and Michael McKenna, eds., *Moral Responsibility and Alternative Possibilities*, (Burlington, VT: Ashgate, 2003).
- Keith Wyma, "Moral Responsibility and Leeway for Action," *American Philosophical Quarterly* 34 (1997): 57-70.

BIOGRAPHICAL SKETCH

Jaclyn Genchi was born in Palm Beach Gardens, Florida, and went to elementary, junior high, and high school in the surrounding areas. She attended Suncoast Community High School, where she received an International Baccalaureate degree. Jaclyn went to Elon College, in Elon, North Carolina, for her first year of undergraduate education, and then transferred to the University of Florida to complete a Bachelor of Arts in philosophy in May 2004. After graduating with her Master of Arts in philosophy in August 2006, she will start law school at Duke Law as a 1L in the graduating class of 2009.