Thesis written by

Adam Shatsky

B.A., Grand Valley State University, 2013

M.A., Kent State University, 2016

Approved by

Michael Byron, Masters Advisor

Deborah Barnbaum, Department Chair, Department of Philosophy

James L. Blank, Dean, College of Arts and Sciences
# TABLE OF CONTENTS

ACKNOWLEDGMENTS ........................................................................................................... v  

INTRODUCTORY SECTION ........................................................................................................... 1  

CHAPTER 1: MORAL RESPONSIBILITY AND THE PRINCIPLE OF ALTERNATIVE POSSIBILITIES

§1.1.1 Moral Responsibility and Its Technical Terms .............................................................. 5  
§1.1.2 The Principle of Alternative Possibilities ...................................................................... 7  
§1.2.1 Frankfurt’s Challenge to PAP .......................................................................................... 10  
§1.2.2 Implications of the Truth of PAP .................................................................................... 12  
§1.2.3 Salient features of Frankfurt-Style Cases ................................................................. 15  
§1.2.4 Genuine Frankfurt-Style Cases ..................................................................................... 17  
§1.2.5 The Dilemma Defense ................................................................................................. 18  
§1.3.1 No-prior-sign Cases .................................................................................................... 19  
§1.3.2 Blockage Cases ........................................................................................................... 21  
§1.3.3 Necessary-condition Cases ......................................................................................... 23  
§1.5 Conclusion ....................................................................................................................... 25  

CHAPTER 2: RECONSIDERING THE FLICKER OF FREEDOM DEFENSE

§2.1 Introduction ..................................................................................................................... 26  
§2.2.1 The shift from traditional to modified Frankfurt-Style Cases ................................. 28  
§2.3.1 Backward-looking Flicker ......................................................................................... 30  
§2.3.2 Forward-looking Flicker ............................................................................................ 32  
§2.3.3 Libertarian Flicker .................................................................................................... 34
Acknowledgements

I have many to thank for this thesis. I would first like to thank my advisor, Dr. Michael Byron. Without his guidance and feedback throughout the contemplating, writing, and editing process, this thesis would have been impossible. I would also like to thank my advisory committee, Dr. Deborah Barnbaum and Dr. Kimberly Garchar, for their intellectual and emotional support and guidance through this process. I owe a great debt to Dr. David Pereplyotchik, whose friendship, guidance, and support through this process has been unparalleled. I am grateful to all members of the Philosophy faculty at both Grand Valley State University and Kent State University for the success of my academic accomplishments. My academic success is also largely due to the support from my family, friends, and lovely girlfriend, Kelsey Mackie. Without their emotional support throughout this process, who knows where I’d be. In particular, I would like to thank my parents, Michelle Harmody and Howard Shatsky, and my step-father, Robert Hooker; without them, I would certainly not be here. Lastly, I would like to dedicate this thesis to Dr. Michael Robinson. Michael’s pedagogy and methodology as an instructor and scholar is an example of what I hope to emulate one day. I have him to thank for both my general passion and specific interests towards philosophy. Here’s to many more nights enjoying a burger, drinks, and great conversations at Brewery Vivant.
Introductory Section

In his article, “Alternate Possibilities and Moral Responsibility,” Harry Frankfurt (1969) famously sparked a new chapter in the contemporary free will debate. To a certain extent, Frankfurt’s contribution parallels Edmund Gettier’s (1963) famous article, “Is Justified True Belief Knowledge?” in the sense that it triggered a vast literature surrounding Frankfurt cases. This thesis in no way attempts to give a comprehensive overview of the literature concerned with Frankfurt cases; I doubt that such a task is possible. In a recent article from a prominent figure in the debate regarding Frankfurt cases, John Martin Fischer (2010) writes:

Yes, the literature is large, intricate, and complicated; I can understand why Harry Frankfurt once complained to me that the literature surrounding his examples was “a young person’s sport.” As I become less qualified to participate in such sports, I seek to find the simple, powerful lesson of the literature—abstracting away from the details. (p. 318; italics my own)

While Fischer’s ultimate aim seems to be figuring out the moral of the story regarding the contemporary debate, I am still young, so I am entering the draft for whatever the “sport” might be: this thesis (only to a certain extent) narrows in on the details of the literature. But, before doing so, I want to acknowledge the broader philosophical context in which my thesis occurs.
First, the larger philosophical issue in my thesis is the freedom condition for moral responsibility. Of course, the issue of moral responsibility is nested in the broader question about the nature of persons. Such a question, however, is outside the scope of this thesis, though it is important to acknowledge the bigger picture of the debate. My aim is not to answer questions concerning the nature of persons, personhood, or what guarantees moral standing, but rather focus on one specific topic within the surrounding philosophical issues. To be more specific, this thesis stems from the following generalization: All morally responsible agents are persons, persons who have achieved a certain sophisticated level of development. In what does this sophisticated level of development consist? Intuitively, it seems that this level allows us to be appropriate candidates for moral praise and blame for our actions—that we may properly be morally responsible, in some sense, for our actions. Attributions of praise and blame seem appropriate only if agents have some control over what they do; we would not typically hold individuals morally responsible for an action over which they had no control (see §1.1.2). In this sense, one way to express the relevant kind of control is to think that, in order for agents to be appropriately called morally responsible for an action, they must be able (or have the power) to do otherwise. This characterization will serve as the starting point of my thesis.

Second, there are two major views concerned with the problem of free will\(^1\): compatibilism and incompatibilism. Compatibilists argue, at minimum, that free will and determinism are compatible (see §1.2.2). Incompatibilists argue, at minimum, that free will

\(^1\) I use ‘free will’ in the most neutral sense possible. I do not intend to subscribe to any theory by this usage. Moreover, for purposes of this thesis, I use ‘will’ and ‘agency’ interchangeably, unless I indicate otherwise.
will and determinism are incompatible (see §1.2.2). Prior to Frankfurt’s article, the free will debate almost exclusively dealt with philosophical issues concerning the compatibility between free will and determinism. Although Frankfurt’s contribution concerns specifically moral responsibility (see §1.1.1) and its relationship to alternative possibilities (see §1.2.1), there are implications that affect both sides of the debate (see §1.2.2). I mention these implications, but it is not the task of this thesis to settle any disputes regarding the broader compatibility issues between free will and determinism.

For the most part, the last two points mention what I will not do, now a word on what I will do. Persons are morally responsible for an action only if they could have done otherwise. This is the Principle of Alternative Possibilities (PAP) (see §1.1.2). Frankfurt offered a putative counterexample that challenged this entailment. That is, he developed a case in which persons can seem intuitively to be morally responsible for an action, and yet they could not have done otherwise (see §1.2.1). Due to his counterexample, some have abandoned PAP and others have sought to defend it. My aim is to defend PAP, but only to a certain extent. I will explain. One line of defense, coined by Fischer (1994), is the flicker of freedom defense. In essence, this defense contends that, even in Frankfurt’s counterexample, agents still maintain the ability to do otherwise and, thus, are morally responsible.

This thesis endorses the flicker defense (see §2.3.1, §2.3.2, §2.3.3, and §2.3.4). Although Fischer both coined the name and developed the defense, he offered a defeater to the flicker defense (see §2.4.1, §2.4.2, and §2.4.3). Fischer’s defeater maintains that the alternatives still available to agents in Frankfurt cases are not significant enough to
ground attributions of moral responsibility. His defeater is the *robustness objection*. As a result of Fischer’s defeater, it is incumbent upon the flicker theorist to defeat his defeater. That is, they must show that the alternatives available to agents in Frankfurt’s case are robust in the sense required in order to properly call them morally responsible for their actions. Strictly speaking, then, the task of this thesis is to defeat a defeater put forth by Fischer. If it happens to turn out that PAP is also true, then all the better for proponents of PAP. But, again, it is not the job of the flicker theorist to preserve the truth of PAP; it is merely to defeat a defeater.

In chapter 3, I provide a brief overview of some of the differing characterizations of robustness (see §3.2.1. and §3.2.2). I will argue that flicker theorists need only respond to Fischer's robustness objection. Flicker theorists that advocate for certain epistemic conditions to robustness seem to miss the mark. I will contend that Michael McKenna (2003), Michael Robinson (2013), and Seth Shabo (2013 and 2015) offer an epistemic condition requirement of robustness that is too strong and implausible (see §3.4.1 and 3.4.2). In this chapter, I will first problematize McKenna and Robinson’s robustness requirement. Lastly, I will return to Seth Shabo’s (2013 and 2015) requirement and argue that his epistemic condition for robustness is too stringent of a condition (see §3.5). As a result, flicker theorists need not engage in the debate over epistemically conditioned robustness requirements.
§1.1.1 Moral Responsibility and Its Technical Terms

It is natural to expect that, since being morally responsible requires the relevant kind of control, and the relevant kind of control here is the ability to do otherwise, we have a (rough) definition of free will. Freedom of the will or free will, however, is a notion at stake in debates about moral responsibility and, as such, to offer a definition at this stage in the discussion would be premature. For now, here is what I take to be a rough characterization of some of the substantive theoretical commitments regarding (see Frankfurt 1969; Pereboom 2003; Robinson 2013) free will: (i) it requires the ability to do otherwise (Robinson's leeway condition); (ii) it requires agents to be the source of their actions (Pereboom’s sourcehood condition); and, (iii) it is required for moral responsibility.

Moral responsibility is responsibility for right and wrong actions; it is simply being an appropriate target for attitudes such as praise and blame. Moreover, it seems appropriate that, in order to be morally responsible, we must have had a choice to take a particular course of action over another. Indeed, we would not (in most cases) hold
persons morally responsible for some action over which they had no control. For example, it would be inappropriate for me to blame my mother for failing to pick me up from school on time because she sustained some injury, which prevented her from being able to (for another example of this sort—namely, an exculpatory reason for being unable to do something—see §1.1.2). In other words, moral responsibility requires that we act freely, per (iii) above. The Principle of Alternative Possibilities (PAP) is the thesis that moral responsibility requires the ability to have done otherwise. It seems, then, that free will and moral responsibility share a necessary condition: the ability to do otherwise. However, discussions of the necessary and sufficient conditions of free will (for the most part) are not the same as ones for moral responsibility. My concern is to examine the ability to do otherwise in regards to moral responsibility.

Only the most recent literature addresses the nature of moral responsibility itself. In a footnote in Frankfurt’s original article, he writes:

The two main concepts employed in the principle of alternate possibilities are “morally responsible” and “could have done otherwise.” To discuss the principle without analyzing either of these concepts may well seem like an attempt at piracy. The reader should take notice that my Jolly Roger is now unfurled. (1969, p. 835)

As Frankfurt rightly claims, such an analysis has been unnecessary for much of the debate concerning issues within free will since both sides seem to have been operating with a similar notion of moral responsibility—namely, a desert-entailing conception of moral responsibility. According to such a conception, persons who are morally
responsible are thereby capable of deserving blame or praise (and punishment or reward) for things they have done.

In light of the foregoing, I intend to engage Frankfurt on his grounds—that is, by presupposing a desert-entailing conception of moral responsibility. Frankfurt assumes a desert-entailing conception of moral responsibility. He makes use of cases of constraint and coercion to motivate his intuition of moral responsibility. Frankfurt argues that, in cases of this sort, the inability to do otherwise plays a crucial role in the explanation for why agents act as they do. And, so I do not plan to say anything more about moral responsibility, as it is broadly construed, because I mean whatever Frankfurt does.

My ultimate aim for this chapter is to offer an overview of some of the literature that is pertinent to this topic. Thus, my goals are threefold. First, I begin with a brief discussion of moral responsibility and its associated technical terms pertinent to this project. Second, I examine the intuitive plausibility of the Principle of Alternative Possibilities and present Harry Frankfurt’s (1969) putative counterexample to it. Third, I discuss the present literature on the Principle of Alternative Possibilities, and I will address three different strategies used to falsify the Principle of Alternative Possibilities.

§1.1.2 The Principle of Alternative Possibilities

Persons are morally responsible for what they have done only if they could have done otherwise. This is the Principle of Alternative Possibilities (PAP). There are several ways to articulate the intuitive idea behind PAP. I will focus on two. Firstly, commonplace in our social practices is the idea that being morally responsible for some
action requires that it is possible for us to not have done that action. PAP expresses the intuitive idea that unless people have a choice about what they do, and unless they can avoid acting as they do, it would be improper to think they are morally responsible for their behavior. For example, if I break my promise to attend your football game, you will likely be predisposed to blame me, unless I have some sort of an adequate excuse. One sort of excuse I might have is that another close friend was in need of immediate medical attention, and I was the only one who could assist him. An excuse of this sort would be known as a justificatory reason for breaking my promise because the duty I have to my friend’s medical needs is more stringent than attending my other friend’s football game. Secondly, another sort of excuse would involve it being out of my control to attend the football game due to some serious injury I sustained, which precluded me from attending the game. This sort of excuse would serve as an exculpatory reason for not keeping my promise. Both of these types of excuses can make me blameless, and thus my friend blaming me for missing his game no longer seems appropriate. Only one of the examples, however, concerns PAP: in the situation where I had a justificatory reason for breaking my promise, I still could have chosen not to do so—there was nothing physically constraining me from fulfilling my promise. An exculpatory reason, then, involves something that is preventing me from doing some action. Furthermore, this example highlights the degree to which being morally responsible, understood in terms of PAP, is firmly embedded in our social practices.

According to David Widerker and Michael McKenna (2003), another way to articulate PAP is to consider the propriety conditions for morally responsible agency.
There are at least two conditions: an epistemic and a freedom condition. Consider the epistemic condition first. If agents are to be morally responsible for what they have done, then they have some understanding (or the ability to understand) of the moral significance of their action:

If Ashley accidentally saves Casey’s life by unwittingly setting down her coffee cup and thereby diverting a stray bullet headed for Casey, fortunate as Casey is, Ashley is not morally responsible for saving Casey’s life in the way she would be had she done so knowingly intending to save Casey. (Widerker and McKenna 2003, p. 2)

I believe Widerker and McKenna rightly contend that this epistemic condition holds some weight in whether agents are morally responsible for what they have done. Their example includes both knowing and intending of an agent, each of which are necessary for Ashley to be praiseworthy. If Ashley intends to save Casey’s life, then she must know that Casey’s life is in danger. In this sense, then, it seems that an essential feature of an intention is some pertinent propositional attitude.

Regarding the freedom condition for morally responsible agency, suppose that Ben suffers an unexpected and uncontrollable heart attack in which, while grasping at his chest, he punches Brian in the face. It seems improper to hold Ben morally responsible—namely, blameworthy—for punching Brian in the face because it was out of his control. Instead, it seems proper for us to hold agents morally responsible only if their actions were done freely, or within their control. Thus, a rather uncontroversial way to express

---

2 It is important to note that these two ways are certainly not the only ways to explain PAP. Philosophers have approached the issue from many different angles.
the freedom condition is in the following principle: agents are morally responsible for what they have done only if they did it freely. Examples similar to Ben’s heart attack confirm this principle, but other examples do so too. Suppose that Brenda is supposed to pick up her daughter, Kelsey, from school. However, on the way to the school Brenda’s car breaks down and requires a tow. The time of the breakdown is such that she cannot rent or borrow another car to pick up Kelsey. As a result, Kelsey’s father, Rob, has to pick Kelsey up from school. We should not hold Brenda blameworthy in this situation. It was out of Brenda’s control to pick up Kelsey from school on the particular day that her car required a tow. Any example that includes an exculpatory reason for why agents cannot do a certain (and perhaps expected) action hinges on the freedom condition. In other words, the freedom condition states that agents must have the right sort of control over their action in order to be held morally responsible for that action. Alternative possibilities (or lack thereof) can explain why we do not hold particular agents morally responsible for certain actions: Ashley didn’t know she was saving Casey’s life; it was out of Ben’s control to not punch Brian in the face; Brenda had no viable alternative to her failure to pick Kelsey up from school. In short, PAP seems so deeply embedded in our social institutions, personal lives, and relationships such that it is intuitively immune to objection(s).

§1.2.1 Frankfurt’s Challenge to PAP

In his seminal 1969 article, “Alternate Possibilities and Moral Responsibility,” Harry Frankfurt challenged the freedom condition for moral responsibility as expressed in PAP.
He offered a case where an agent is morally responsible for some action, even though he could not have done otherwise. Frankfurt’s original thought-experiment is as follows:

Suppose someone—Black, let us say—wants Jones to perform a certain action. Black is prepared to go to considerable lengths to get his way, but he prefers to avoid showing his hand unnecessarily. So he waits until Jones is about to make up his mind what to do, and does nothing unless it is clear to him (Black is an excellent judge of such things) that Jones is going to decide to do something other than what he wants him to do. If it does become clear that Jones is going to decide to do something else, Black takes effective steps to ensure that Jones decides to do, and that he does do, what he wants him to do. Whatever Jones’s initial preferences are inclinations, then, Black will have his way. (Frankfurt 1969, p. 835)

As it turns out, however, Black does not need to intervene because Jones’s action is exactly what Black wants. So, what this thought experiment putatively shows is that Jones is morally responsible for some action A, yet he could not have done otherwise than action A. In the quoted passage, Frankfurt does not mention the action for which Jones is morally responsible, which is killing a third man, Smith. If Jones had decided not to kill Smith, Black would have intervened. However, since Jones decided to kill Smith, Black did not have to intervene. Consequently, Frankfurt’s case is apparently successful in two respects. First, Jones chose to kill Smith. Nothing coerced him into that choice, he was able to act on that choice, and he succeeded at being the proximate cause of Smith’s death—Jones seems to be morally responsible. Second, the setup of
Frankfurt’s case supports the following two subjunctive conditionals: ‘If Jones were to decide to shoot Smith, he would shoot Smith’ and ‘If Jones were to decide not to shoot Smith, he would shoot Smith’. That is, no matter what Jones decides, he will shoot Smith; he is unable to do otherwise than kill Smith. Thus, Frankfurt offers a putative counterexample to PAP, which challenges the freedom condition for moral responsibility (as expressed in PAP). In other words, moral responsibility seems not to presuppose alternative possibilities, so PAP seems false.

§1.2.2 Implications of the Truth of PAP

If Frankfurt’s argument is sound, the implications are significant. These implications, though not all within the scope of this project, exist within the broader debate concerning the relationship between free will and determinism. Determinism is the thesis that every event is causally determined. More specifically, it is the thesis that all events are causally determined, in the sense that the laws of nature $n$, in conjunction with a description of the state of affairs at time $t$, will entail the occurrence of every event after $t$.³ Compatibilists maintain that free will and moral responsibility are compatible with determinism, whereas incompatibilists hold that they are not. Some incompatibilists additionally argue that determinism is false and we are in fact free and morally responsible for our actions. These incompatibilists are known as libertarians.

There are at least four implications that would follow from the soundness of Frankfurt’s argument. First, if Frankfurt’s argument is sound, then it would affect both

---

³ This thesis also is backwards working in that, if true, it can explain all events prior to time $t$. 
compatibilists and incompatibilists’ understanding of precisely what sort of freedom is required for moral responsibility; prior to Frankfurt’s article (for the most part) both parties accepted PAP (Widerker and McKenna 2003, p. 5). But, if PAP is false, then an account of freedom is necessary. Both compatibilists and incompatibilists would be required to reexamine their notions of freedom. Even though we know that Frankfurt challenged the freedom condition as it pertains to moral responsibility, the challenge remains as to how to express freedom that does not involve alternative possibilities.

The second implication of the success of Frankfurt’s argument is that it would challenge a classic argument for the incompatibility of freedom and determinism (Widerker and McKenna 2003, p. 6). Such an argument hinges on PAP:

1. Agents are morally responsible for what they have done only if they could have done otherwise (PAP).

2. If determinism is true, no agent can do otherwise.

3. Thus, if determinism is true, agents are not morally responsible for they have done.

By undermining premise (1), Frankfurt’s argument puts the burden on incompatibilists and libertarians to account for why, provided that determinism is true, agents are not morally responsible (Widerker and McKenna 2003, p. 6). This burden does not threaten the compatibilist thesis (broadly understood), since in principle they accept the compatibility between determinism and the ability to do otherwise.

The third implication concerns compatibilists. Since the truth of PAP seems consistent with libertarian intuitions, it is easy to see how determinism comes into
conflict with PAP’s requirement of alternative possibilities. At the time the dialectic was framed, compatibilists had to defend a set of substantive theoretical commitments that were nearly untenable. However, the abandonment of PAP simultaneously entails the abandonment of this untenable position (Widerker and McKenna 2003, p. 6). It seems that they are no longer in this dialectically subordinate position because, “when PAP was assumed, intuition spoke on behalf of the incompatibilists”—it is intuitive to see the conflict between there being only one unique future (i.e., determinism) and PAP’s requirement for alternative possibilities (Widerker and McKenna 2003, p. 6).

The fourth implication is that, without PAP, a compatibilist can grant premise (2) (if determinism is true, no agent can do otherwise) in the aforementioned argument for incompatibilism. There are conceptions of control that are compatible with determinism and do not require the ability to do otherwise. What I mean by this is that compatibilists can accept that determinism is incompatible with the freedom to do otherwise, and yet deny the conclusion. This form of compatibilism is known as semi-compatibilism.\(^4\) Semi-compatibilism maintains that determinism is incompatible with the ability to do otherwise, but there is compatibility between determinism and the freedom required for moral responsibility. Accordingly, such an account of freedom would not consist of the ability to do otherwise (Widerker and McKenna 2003, p. 6).

\(^4\) The most promising argument for semi-compatibilism can be found in John Martin Fischer’s *The Metaphysics of Free Will*, ch. 8, pp. 160-190.
§1.2.3 Salient features of Frankfurt-style Cases

There are at least six distinct features on which Frankfurt’s case (and Frankfurt-style cases) rely. First, there are certain alternative possibilities that are available to Jones. The example works only under the assumption that Jones’s body will or will not demonstrate a certain sign. These signs are themselves alternatives in the sense that they are different ways in which the future might unfold. Indeed, in Frankfurt’s case Jones killing Smith is unavoidable, but there is more than way in which it could unfold: either by Black’s intervention or out of Jones’s own accord. These alternatives, though quite small, are built into the example. Thus, there still remains a limited sense in which Jones could have done otherwise. John Martin Fischer calls these flickers of freedom. Fischer means that there are alternative possibilities in Frankfurt’s case just as there are in ordinary situations (where there is no counterfactual intervener). In an ordinary situation, one that does not include the counterfactual intervener, Jones could have decided not to kill Smith and not done so. Even with Black in the picture, Jones can still decide not to kill Smith. It is just that, if he decides not to kill Smith, Black will intervene. What Fischer highlights is that, in Frankfurt’s case, there is still an alternative possibility available to Jones—namely, deciding not to kill Smith—and this alternative shows a very small way—merely a flicker—in which Jones is free to have done otherwise.

---

5 The term ‘flickers of freedom’ was coined by Fischer (1994) to refer to the alternatives still available to agents in Frankfurt-style cases. I discuss this term in-depth in chapters 2 and 3.

6 By “counterfactual intervener,” I mean someone or something (whether it be an individual, like Black, lurking in the background with sinister intent, or a some sort of neurological device) that, had Jones decided to do something other than what this someone or something desires, this someone or something would have intervened.
in the sense that he could have decided not to do the action that Black wants him to do. His decision is free, even though his action is not.

The second important feature of Frankfurt’s case is that there is a sign, e.g., a blush, twitch, or neurological pattern that indicates an agent’s future action. This sign, which is called a prior sign, occurs before an agent acts. Accordingly, examples involving such a sign are often known as prior sign examples (Widerker and McKenna 2003, p. 7).

The third important feature of Frankfurt cases is that this sign is supposed to play an essential role of subsequent free action and, as such, these signs are considered reliable indicators of what will obtain (Widerker and McKenna 2003, p. 7). In other words, the sign the agent demonstrates in these cases is a reliable indicator of what the agent will do.

The first acts that are in the voluntary control of agents are the loci of free agency and moral responsibility. The fourth important feature concerns these acts. Simple mental acts such as choosing, deciding, forming intentions, etc. are integral to these cases (Widerker and McKenna 2003, p. 7).

The fifth important feature is that Frankfurt cases are said to involve an ensuring condition (Widerker and McKenna 2003, p. 7). This term refers to the entire setup of these cases involving a counterfactual intervener. A condition that ensures the unavoidability of the agents’ action or the outcome itself is essential to these cases.

The sixth, and last, salient feature of Frankfurt cases involves a sign that triggers intervention by the means of a fail-safe mechanism. Call this sign a triggering event
In other words, in the counterfactual scenario, it ‘activates’ the intervening mechanism in what would have been an otherwise freely willed course of action.\footnote{One might rightly question the difference between a triggering event and a prior sign. The prior sign is an event that causes intervention before the act. And, while the triggering event is still a sign that causes intervention, it occurs during the action. Other than this, however, there does not seem to be much difference.}

\section*{§1.2.4 Genuine Frankfurt-Style Cases}

There are now a number of variations of Frankfurt’s case, known as Frankfurt-style cases, which intend to do the same work as Frankfurt’s original case. As in Frankfurt’s case, the aim of Frankfurt-style cases is to undermine PAP, and the most recent cases attempt to do so by presenting situations in which an agent is in a \textit{genuine Frankfurt-style case} (GFSC), which can be characterized in the following way:\footnote{Widerker and McKenna refer to GFSC as an “IRR-situation” (see p. 3, 2003). The usage of “IRR” is not limited to Widerker and McKenna, many others use this abbreviation. Although many use “IRR,” they do so without any explanation of what it means. So, I am going with my own version, which seems more appropriate, though the two can be used interchangeably.}

There may be circumstances in which a person performs some action, which, although they make it impossible for him to avoid performing that action, they in no way bring it about that he performs it. (Widerker and McKenna 2003, p. 3)

Frankfurt’s case describes an agent in GFSC, but in no way do these ‘circumstances’ \textit{cause} Jones to act in any way; they merely make it so Jones could not have done otherwise than what he in fact did. The success of Frankfurt examples hinges on whether they genuinely describe GFSC. Some maintain that Frankfurt examples satisfy GFSC, while others deny it.
§1.2.5 The Dilemma Defense

Due to the conditional nature of PAP, a threat to it must be a case that includes a true antecedent and a false consequent. Frankfurt must show a case of morally responsible action and no alternative possibilities available. Defenders of PAP must show one of two things. Either (i) Jones’s sign, or whatever causes him to twitch, is causally sufficient for his subsequent action. In this situation, then, he would not be morally responsible: the same thing that makes it impossible for Jones to avoid killing Smith (that is, the process that leads to the twitch) is also what causes him to kill Smith and so the prior sign presupposes a deterministic relation between an agent’s free action and its causal antecedents. Or, (ii) Jones’s sign, or whatever causes it, is not causally sufficient for Jones killing Smith, but merely accurately predicts what he will do. But, now it seems that, if Jones demonstrates some sign (and Black does not intervene), he still retains the possibility of doing otherwise.

In short, in (i), we do not have a case that is a GFSC because the circumstances cause Jones to kill Smith. The presupposition that there is a deterministic relationship between the locus of free action and its causal antecedent begs the question against incompatibilists. So, incompatibilists will deny that this is a case of morally responsible action—the case poses no threat to PAP. This lemma denies the truth of the antecedent. And, in (ii), we do not have a case that is a GFSC because Jones retains alternative possibilities: it is merely probable that Jones will kill Smith, even if Black intervenes. So again, the case poses no threat to PAP. This lemma defends the truth of the consequent.
If either the antecedent is true or the consequent is false, then PAP is true. This line of defense is the dilemma defense of PAP. As a result of the dilemma defense, Frankfurt-style cases are only successful given determinism, which is a view most defenders of PAP reject. We now turn to such cases.

§1.3.1 No-prior-sign Cases

One such case, developed by Alfred Mele and David Robb, argues that there need not be any prior sign present in order to constitute a GFSC. Their case is as follows:

Our scenario features an agent, Bob, who inhabits a world at which determinism is false … 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 \( t_2 \)) … 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. (1998, pp. 101-102)

Mele and Robb’s case makes use of an ensuring condition that is not dependent on any sign. Such a condition is designed to bring about the desired action “only in the counterfactual event that the agent does not freely perform that action” (Widerker and McKenna 2003, p. 9). Of course, Bob performs the action on his own without its being causally determined. This example can be explained by considering two distinct processes—one that is deterministic and one that is not—both of which aim at a certain outcome. The deterministic process is designed to bring “about the outcome under all circumstances except the one in which the other process brings about the outcome” (Widerker and McKenna 2003, p. 9). The indeterministic process, on the other hand, involves an free action. According to Widerker and McKenna:

Suppose that, in the actual world, the indeterministic process issues in the desired action. But imagine that, if the indeterministic process were to fail to do so, the deterministic process would. Neither process, however, involves any prior sign of the action, and both are independent of each other. All that happens is that the indeterministic process trumps the deterministic process at the moment of action. (2003, p. 9)

What is crucial to this case is that agents acts on their own in some indeterministic manner. But, even though agents may act on their own via an indeterministic process, there is no alternative available to them due to the independent deterministic causal chain.
§1.3.2 Blockage Cases

Michael McKenna (2003) has introduced another class of cases that do not rely upon a prior sign, known as “limited-blockage cases,” which he calls Brain Malfunction. He writes:

The mild-mannered philosophy professor Casper comes upon a completely unexpected and highly unusual opportunity. He has just entered a room and is standing in front of a technologically state of the art ‘Make-it-the-Case Device’. Assume that Casper is justified in his true belief that the Make-it-the-Case Device is reliable and not merely a hoax.

On a large television screen at the top of the Make-it-the-Case Device appears a man dressed as a genie. The genie speaks:

Casper, just beneath this screen are two buttons, one marked ‘The Morally Good Thing to Do’ and another marked ‘The Morally Bad Thing to Do’. Let us abbreviate them as ‘Good’ and ‘Bad’ respectively. If you press the Bad button you will immediately make it the case that one million dollars are deposited into your bank account. The money will be drawn, in one-dollar increments, from the savings accounts of one million college professors. The transaction will be untraceable. If you press the Good button you will immediately make it that case that an entire village of people in the Amazon is cured of an otherwise fatal disease. Saving the villagers will not involve any money and by doing so you will not be stealing from your peers. You cannot select both buttons and this
opportunity will not present itself again. You will have ten seconds to select your option.

A timer appears on the screen and begins to count down from ten.

Casper pauses to consider these two options, quickly assessing the import of each. He considers the article he read in last Sunday’s *New York Times* on the villagers’ plight. He is fully aware of the urgency of their condition. He also considers his sparse salary as a philosophy professor, and he squirms at the thought of stealing from his peers. Imagining that shiny red Mercedes convertible roadster in the window, as the counter ticks away from 3 seconds to 2 greedily he takes the plunge and presses the Bad button. ‘Ah, dinner out tonight!’ Casper thinks to himself.

As it turns out, although Casper was unaware of this difficulty, and although there is no reason Casper should have been aware of this difficulty, at the time at which Casper greedily decided to press the Bad button, Casper had a small lesion on his brain that blocked the neural pathway constitutive of (or correlated with) a decision to push the Good button during that ten second interval. Casper could not have decided to press the Good button.

The lesion on Casper’s brain plays no causal role in his decision-making, and he acts just as he would if the lesion had not been present. Nonetheless, it seems reasonable enough to think Casper is morally responsible for what he does.

Further notice that blockage examples, like Brain Malfunction, rely on no prior sign of an agent’s action. Rather, they involve a blockage of any path other than the
desired action, and this blockage is in no way causally related to the agent choosing and performing the desired action. The construction of blockage cases is not limited to Brain Malfunction (see Hunt, 2000).

§1.3.3 Necessary-condition Cases

The last class of cases I will mention in this chapter is known as necessary-condition cases. Such cases attempt to revitalize the use of prior-sign examples. So, in this respect, they differ drastically from blockage and no-prior-sign examples. The most recognized case of this sort comes from Derk Pereboom (2000), which he calls “Tax Evasion”:

Tax Evasion: 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 that in this situation he could fail to choose to evade taxes is for moral reasons. (The phrase “failing to choose to evade taxes” is meant to encompass not choosing to evade taxes and choosing not to evade taxes.) 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 choose 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.

As Pereboom has demonstrated in Tax Evasion, the use of a prior sign does not presuppose a causally sufficient condition for a subsequent free action. Rather, the prior sign simply flags the presence of a condition that is strictly necessary for any alternative to the desired course of action. In the world Joe inhabits, he freely wills a course of action, and yet he could not have done otherwise. For, if Joe were merely to satisfy some necessary—but not sufficient—condition for doing otherwise, then an intervener would ensure a certain course of action, in which case Joe would have still evaded taxes. The necessary condition, which in Joe’s case is attaining a higher level of attentiveness, must involve some specific kind of reason for acting otherwise. In short, Pereboom argues that Tax Evasion represents a case wherein an agent demonstrates a prior sign, though not of the illicit type, is still morally responsible for said action, and yet could not have done otherwise.
§1.4 Conclusion

In this chapter I offered a brief analysis of what it is to be morally responsible. I then honed in on what was once a widely accepted necessary condition of moral responsibility—namely, PAP. Since Frankfurt’s seminal work, many philosophers have discussed Frankfurt-style cases and the implications of their soundness for PAP. As the debate has developed, proponents of Frankfurt-style cases have been forced to offer cases which avoid prior signs of the illicit type and which are not guilty of begging the question against incompatibilists. The worry for proponents of PAP, on the other hand, has persisted such that they are still left with the same task: demonstrate that Frankfurt-style cases do not satisfy GFSC. The means by which those who endorse PAP do this is dialectically open. I believe one promising way to maintain the truth of PAP lies in the flicker defense strategy.

As I argue in chapter 2, proponents of PAP have available to them a strategy known as the flicker of freedom defense. Flicker theorists’ aim is to preserve the truth of PAP by arguing that, despite these Frankfurt-style cases, there is still a sense in which agents could have done otherwise—namely, in the sense that they could have decided other than what they in fact did. Such a decision is supposed to lay the groundwork for a defense of PAP against Frankfurt-style cases.
Chapter 2

Reconsidering the Flicker of Freedom Defense

§2.1 Introduction

According to traditional arguments for incompatibilism, we are not morally responsible for what we cannot avoid doing, and we cannot avoid doing what we are deterministically caused to do. Compatibilists traditionally respond by arguing that causally determined actions meet the relevant avoidability requirement (i.e., PAP). As discussed in chapter 1 (§1.3.1), by challenging PAP itself, Frankfurt opened a new chapter in the free-will debate. Frankfurt offered a case that purports to be a counterexample to PAP. Whether Frankfurt falsified PAP has been a matter of controversy. In particular, the success of Frankfurt’s actual case, which fundamentally involves a counterfactual intervener to guarantee that an agent will act in a certain way, hinges on the existence of a perfectly reliable indicator (i.e., prior sign) of what an agent will freely do if left to act on his own. The relationship between these prior signs and what an agent does, however, seems to be a deterministic one. That is, these indicators seem to presuppose a causally deterministic relation between an agent’s free action and its causal antecedents, which (in principle) incompatibilists reject because it precludes responsibility. As a result of this objection, there has been considerable effort to
construct Frankfurt-style cases that do not involve a deterministic prior sign (see §1.2.5), thereby restoring the force of the counterexample to PAP.

According to Michael Robinson (2012), one consequence of the desire to reformulate Frankfurt-style cases is that doing so allows for the revival of the so-called *flicker of freedom* defense strategy of alternative possibilities. The ultimate aim of flicker theorists is to show that, even though Frankfurt-style cases succeed in eliminating many alternative possibilities, there are still some alternatives—namely, the flickers of freedom—that remain.

According to John Martin Fischer (1994), there are at least four different versions of the flicker defense. The versions differ in how they identify and conceive of the flicker of freedom in Frankfurt-style cases. Before I discuss the specifics of each version, a brief account of the significance of a flicker of freedom is in order. Recall that one condition Frankfurt-style cases must satisfy is that there be some alternatives available to an agent; if there were not any, then Frankfurt-style cases would be guilty of begging the question against incompatibilists. One flicker strategy, which I will discuss in Chapter 3 (§3.3.1 and §3.3.2), argues that, if PAP is a thesis about direct (i.e., non-derivative) responsibility, then flicker theorists seem to have a successful response to modified Frankfurt-style cases.

In this chapter I briefly summarize the shift from traditional to modified Frankfurt-style counterexamples. Then, I discuss four different versions of the flicker defense, as introduced by John Martin Fischer. I will present Fischer’s defeater to the flicker defense. Lastly, since *traditional* Frankfurt-style cases are guilty of begging the
question against incompatibilists, I evaluate which versions of the flicker defense strategy are successful regarding *modified* Frankfurt-style cases.

§2.2.1 The Shift From Traditional to Modified Frankfurt-Style Cases

Michael Robinson (2012) introduces the following traditional Frankfurt-style case:

Jones is a disgruntled postal worker who goes to work on Monday morning unaware that his boss intends to fire him just before noon. Black, an evil neuroscientist, outraged by the increased cost of domestic postage and aware of what Jones’s boss plans to do, hopes that getting fired will make Jones so angry that he decides to burn the place to the ground. Black would prefer that Jones do this on his own. Not wanting to take any chances, however, Black has, unbeknownst to Jones, implanted a special device in Jones’s brain which will enable him to cause Jones to decide to set fire to the post office (and do so) in the event that Jones gives any indication he will not do so on his own. If just before noon Black (somehow) is able to tell that Jones is not going to set fire to the post office on his own, Black will activate the device via remote control, sending electrical impulses to specific parts of Jones’s brain, causing him to decide to set fire to the post office and then do so. As it happens, upon being fired, and without any interference from Black, Jones flips out, punches his boss in the face, and
burns the place to the ground. In short, Jones goes postal. Let us call this case “Post Office.” (Robinson, pp. 178–179)

As previously discussed, the problem with these kinds of cases is the way in which Black is able to know what Jones will freely do if left to act on his own. In other words, cases of this sort presuppose a deterministic relation between Jones’s action and its causal antecedents.

In order to rectify this worry, some philosophers have attempted to get rid of prior signs in Frankfurt-style cases. What is largely considered the most influential attempt in this vein is from Mele and Robb (1998 and 2003). They propose to modifying the cases in the following way:

[S]uppose that Jones’s brain operates in an indeterministic manner. Further suppose that instead of waiting to see whether Jones gives some indication he will set fire to the post office on his own, Black implants a deterministic device in Jones’s brain that at noon will issue in a decision to set ablaze unless Jones decides on his own to set the fire. In the event that, at noon, Jones’s own cognitive faculties issue in a decision to set fire to the post office, the deterministic mechanism implanted by Black will at that very moment be causally preempted from issuing in a decision. The significance of the preemption’s occurring precisely at noon is that it guarantees that at no point leading up to and including noon is it possible for Jones to avoid deciding to set the fire. Furthermore, since

---

10 Readers will notice that this case is nearly identical to Frankfurt’s case with only a few differences. One crucial difference, which will be evident soon enough, is that this case is construed with a special emphasis on the phrase “on his own” (in reference to Jones). This is part of one version of the flicker defense—namely, the “fine-grained” flicker strategy.
this deterministic mechanism is screened off from Jones’s consciousness and plays no causal role in his deciding to act as he does, he looks to be genuinely morally responsible for his actions. Call this “Modified Post Office.”\textsuperscript{11} (Robinson, p. 180)

Like Robinson, I do not wish to engage in the controversy over “the theoretical possibility of the kind of occurrent causal preemption (or ‘trumping’) to which it appeals” (p. 181). There will always be difficulties in offering a case where the ensuring condition guarantees that an agent could not have done otherwise and yet plays no causal role in the agent’s action. That said, the point of including it here is not only as a point of reference for the reader in this chapter—for, if there is any confusion on the differences between traditional and modified Frankfurt-style cases, the reader may reference this section—but also to show that even modified Frankfurt-style cases, which attempt to get rid of prior signs, are not immune to (certain versions of) the flicker defense.

\textbf{§2.3.1 Backward-looking flicker}

The first flicker defense strategy involves tracing “backward” along the alternative sequence of events. In Post Office, Jones deliberates in the normal fashion, decides to set fire to the post office, and does so on his own. He is neither able to decide to do otherwise nor can he do otherwise. Suppose the alternative scenario were true, however: Jones \textit{begins} to decide to not set the fire. If this alternative scenario occurs, Black would

\textsuperscript{11} Mele and Robb’s case is not without problems of its own. The primary controversy over cases in this vein is with “the theoretical possibility of the kind of occurrent causal preemption (or ‘trumping’) to which it appeals” (Robinson, p. 181). For a detailed response to these considerations, as well as a detailed response to the specific kind of mechanism they have in mind, see Mele and Robb (2003).
immediately intervene; we can say, however, that Jones begins to decide to not start the fire. Thus, Jones, in the very least, has the power to initiate, though not complete, the decision to do otherwise. It is in this sense, then, that there is a flicker of freedom.

One can further imagine a case in which Jones is predisposed to show some sign that is a reliable indicator for his subsequent behavior prior to his beginning to form the intention. For example, suppose that Jones would blush red prior to initiating any decision-making process if and only if he was not going to start the fire. If Jones shows this blush, then Black could prevent him in making the relevant decision (Blumenfeld, 1971).

Herein, again, lies a flicker. Even when Jones has the power to show the relevant sign—a blush or neurophysiological pattern that is indicative of what the action will be—there is a flicker. It is difficult to imagine a Frankfurt-style case that has absolutely no such flicker. As Fischer points out, “a Frankfurt-style case must have an alternative sequence in which intervention is triggered in some fashion or other, and it is hard to see how to avoid the idea that the triggering event can serve as the flicker of freedom” (Fischer, p. 31). Therefore, it is at least intuitively plausible that, no matter the sophistication of the case, if one traces from the event caused by the agent to the locus of action, one will find a flicker of freedom.
§2.3.2 Forward-looking flicker

Similar to the Backward-looking flicker, there is a flicker strategy that involves tracing in the opposite way along the alternative sequence.\(^{12}\) Such an approach is “forward” looking in the sense that it proceeds from the agent to the event caused by the agent, until one arrives at the end point. At that point, there will be a flicker of freedom. This strategy is based on the following assumption: when agents perform an action, they cause some concrete particular event to obtain. Also, when agents are morally responsible for performing a particular act, they are thereby morally responsible for causing the relevant concrete event to obtain.

According to Fischer, how this version differs from the backward-looking flicker is that this strategy is committed to a controversial essentialist principle of event-individuation (van Inwagen, 1983). Like van Inwagen (1983), Fischer endorses the strongest version of this principle (see Fischer, p. 32).

According to Fischer, the most robust version of this principle is as follows: “all actual causal antecedents of a particular event are essential to it; thus, if a given event \(e\) occurs in the actual world, then any possible event with any different causal antecedent would not be identical to \(e\)” (p. 32). To put this in the context of Post Office, recall how the example unfolded. Jones chose to set fire to the post office on his own, without any intervention from Black. Counterfactually, however, had Jones demonstrated some sign—namely, a decision to not start the fire—that subsequently led to Black’s intervention, Jones would have still set fire to the post office but only because of Black’s

\(^{12}\) Elsewhere in relevant literature this defense is known as the “act-individuation” flicker strategy, which was first introduced by van Inwagen (1978). For a more current defense of this version, see Pettit (2005).
intervention. Under the counterfactual scenario, Jones would indeed have set the post office ablaze. As things actually happened, Black played no role in the causal background of the event. In the counterfactual scenario, though, Black does play a causal role. Thus, according to the essentialist principle, the actual event cannot be identical to the hypothetical event (i.e., the counterfactual event of Jones deciding not to start the fire). This implies that Jones could have caused a distinct particular event to occur. Herein lies the flicker of freedom. Although Jones cannot bring it about that he does not start the fire, he does possess the power to bring about a different event-particular.

Insofar as responsibility for an action entails (among other things) responsibility for bringing about a concrete (i.e., actual) event-particular, then “responsibility for action involves alternative possibilities, even in Frankfurt-style cases” (Fischer, p. 32).

If one wishes to defend the forward-looking version of the flicker of freedom strategy, one must distinguish the notions of bringing about a particular concrete event and bringing about an event of a general type. Jones cannot avoid the event of the general type—that is, he cannot avoid setting fire to the post office. What he can avoid is bringing about the particular concrete event that subsequently leads to starting the fire. This view suggests that Frankfurt-style cases—which sought to show that agents can be morally responsible for an event even though they cannot avoid bringing it about—rely on some illicit support from insufficiently distinguishing between bringing about a particular concrete event and a concrete event of a general type (see van Inwagen, 1983). Whatever the case, the idea behind this strategy hinges on the consequences of distinguishing between different particular events. Furthermore, van Inwagen’s point is
that there is a distinction among an *act-type* and an *act-token*. Act-types, which are events of general type, are actions of a general sort. Act-tokens, on the other hand, are particular concrete events, which are specific instances of an action. For example, the act of driving to work is an act type, whereas the particular way that I drove to Kent State University this week is an act token. Van Inwagen’s strategy, then, by appealing to this distinction is that we are responsible for act-tokens, not types. That is not to say, however, we are not responsible for act-types; it is simply that our being responsible for an act-type depends on our being responsible for its associated act-token. It is incumbent upon the proponent of this strategy to answer vexed metaphysical questions of what it is that individuates actions.

§2.3.3 Libertarian flicker

Another strategy one may endorse to respond to Frankfurt-style cases is associated with the libertarian picture of agency (see Fischer, p. 33). I will discuss one potential libertarian response, but libertarian responses are by no means limited to the following.

According to some libertarians, that which distinguishes an action from a mere event is that an action is preceded by a *volition* and that this volition is *agent-caused*. Agent-causation is the kind of causation that is not reducible to any form of event-causation. These libertarians assume that when an agent causes a volition through this

---

13 The analogy to act ‘types’ and ‘tokens’ is not without its own controversy. Joshua Spencer (2013), for instance, argues that, the relationship between driving to work and the particular way that I drove to Kent State University, “is more like the relationship of a determinable to a determinate rather than a type to a token” (Spencer, 150, footnote 1). But, whether Spencer is correct is outside the scope of this thesis; nothing of importance hinges on exactly which relationship obtains.

14 Elsewhere in the literature this has been referred to as the “agent-causation” strategy.
sort of causation—namely, agent-causation—nothing causes the agent to perform this volition. That is, there is no external factor playing a causal role in an agent’s volition. Distinguishing between an action from an event by itself does not draw a sufficient picture of libertarian agency. The additional clarification for agent-causation as the sort of causation endorsed is required for (some) libertarian accounts of agency.

To put this strategy to use, one must adopt this picture of agency. Recall, again, Post Office. As it actually turns out, we can say that Jones agent-causes his volition to set fire to the post office. That is, Jones deliberates, decides, and acts on his own in starting the fire. In the counterfactual scenario, however, it is stipulated that if Jones were to show any sign such that he might not start the fire, Black would intervene and neurologically ensure that he start the fire. In such a scenario, we can grant that Jones had some mental state that consisted of a decision to start the fire, but he would not have agent-caused a volition to initiate the fire because his volition was caused by some external factor—namely, Black. Thus, according to this libertarian picture, Jones maintains the power to refrain from agent-causing his volition to set fire to the post office. Herein lies Jones’s flicker: although he actually agent-causes a volition to set fire to the post office, he also has it in his power not to agent-cause this volition. And, since PAP is a thesis that pertains to both actions and omissions, there is still a sense in which Jones could have done otherwise (even in Frankfurt-style cases).
§2.3.4 Fine-grained Flicker

So far, all of the aforementioned flicker strategies look carefully at the counterfactual sequence and, from there, generate an alternative possibility that might otherwise have gone unnoticed. The last strategy differs from the foregoing pictures in that it begins with a careful look at the actual sequence. Such a strategy strives to answer the following question: for what is the agent (putatively) morally responsible? Akin to the foregoing strategies, answering this question will bring to the fore that there are in fact alternative possibilities available (see Naylor, 1984; Stump, 1996; Robinson, 2012; Chapter 3, especially §3.3.1 and §3.3.2).

The point can be put as follows. Someone might claim that what we ‘actually’ hold Jones morally responsible for is not ‘setting fire to the post office,’ or even ‘deciding to set fire to the post office.’ Instead, we hold Jones morally responsible for ‘setting fire to the post office on his own,’ or ‘deciding to set fire to the post office on his own.’ “On his own” means “not as a result of some intervention (such as Black).” If this is in fact what we hold Jones morally responsible for, then there clearly is an alternative possibility. That is, simply ‘setting fire to the post office’ is ambiguous as to whether it refers to the situation in which Jones sets fire to the post office on his own or as the result of Black.

§2.4.1 Fischer’s Defeater: Robustness Objection

Despite its initial plausibility, the flicker of freedom agreement is not the most attractive view in the literature. Fischer (1982, 1994, and 1999), ironically, is widely regarded as
offering the first and most influential rebuttal to flicker theorists.\textsuperscript{15} Fischer’s primary concern with flicker theorists is his notion of \textit{robustness}. As he writes,

\begin{quote}
[M]y basic worry is that this alternative possibility is not sufficiently robust to ground the relevant attributions of moral responsibility. Put in other words, even if the possible event at the terminus of the alternative sequence (in the case of Jones and Black) is indeed an alternative possibility, it is highly implausible to suppose that it is \textit{in virtue} of the existence of such an alternative possibility that Jones is morally responsible for what he does. I suggest that it is not enough for the flicker theorist to analyze the relevant range of cases in such a way as to identify an alternative possibility. Although this is surely the first step, it is not enough to establish the flicker-of-freedom view, because what needs also to be shown is that these alternative possibilities \textit{play a certain role} in the appropriate understanding of the cases. That is, it needs to be shown that these alternative possibilities \textit{ground} our attributions of moral responsibility. And this is what I find puzzling and implausible. (Fischer, pp. 34-35)
\end{quote}

Fischer’s point can be put as follows: the basic picture of control that underlies PAP is a garden of forking paths. By that, I mean that the future includes various paths that are not yet determined, and agents have the power to choose which path they would like to obtain. In other words, agents can, at various points in their life, choose which genuinely

\textsuperscript{15} Fischer’s defeater is specifically addressing the forward-looking flicker; however, Fischer does not believe that he has a decisive rebuttal to the strategy, but rather a series of considerations that count against the strategy. In a series of work, Fischer attempts to establish what is now known as an \textit{adequacy condition} between moral responsibility and the right kind of alternative possibilities for action. For summaries of the “adequacy condition” (and refined versions of it) see Pereboom 2001, p. 26; 2003, p. 188; 2012, p. 301. See also Shabo 2013 and 2015, and McKenna 2003, p. 208.
open pathway they would like to take. For example, it is possible that I can either continue typing or take a ten minute coffee break … Now that I’ve finished my coffee: the existence of these merely possible open pathways is supposed to be integral to the relevant kind of control. Fischer contends, however, that it is not in virtue of the possibility available (whether to take a ten minute coffee break or continue typing that shows I had the relevant kind of control). As Fischer writes, “How exactly could the existence of various alternative pathways along which the agent does not act freely render it true that the agent has the relevant kind of control?” (p. 35). In other words, Fischer is challenging the explanatory role of the existence of alternative possibilities qua possibilities as they pertain to attributions of moral responsibility. According to Fischer, then, the flicker theorists’ strategy seems to hinge on an assumption: the existence of agents’ available alternative possibilities (at least in part) grounds agents’ moral responsibility.

§2.4.2 Objection to Libertarian flicker

Fischer believes that parallel worries confront other versions. I turn now to his critique of the libertarian flicker strategy. Recall that, in order for agents to be morally responsible, this approach requires that agents have the power not to cause the volition that they actually cause. It is true in Post Office that Jones has the power not to cause his volition to start the fire. Fischer contends that, “even so, in the alternative sequence Jones does not form an intention to refrain from causing the volition in question [the volition to start the fire] and then proceed to carry out this intention in an appropriate way” (p. 36).
According to Fischer, one may grant that Jones has the power not to cause a volition to start the fire. He could still have the volition, but (in the counterfactual sequence) it would not be one that he causes. The nature of Black’s intervention is such that it would not be true that Jones freely refrains from causing the volition to start the fire. Fischer thus claims that, even if there is some flicker of freedom at work here, it does not “seem capable of playing the requisite role in grounding ascriptions of moral responsibility—it does not seem sufficiently robust” (Fischer, p. 37).

§2.4.3 Objection to Fine-Grained Version

Recall that the fine-grained strategy contends that what Jones is really morally responsible for is ‘setting fire to the post office on his own.’ If Jones were to merely ‘start the fire’ (or if this were what he was morally responsible for), it could be argued that he was forced, or coerced, by Black to do the deed. But, again, Fischer retorts that this alternative possibility lacks robustness because there is no sense in which Jones can do anything other than start the fire. In the alternative sequence, Jones does not freely refrain from ‘setting fire to the post office on his own.’ Further, Jones does not deliberate and choose the possibility to not start the fire on his own, this is rather done by Black’s intervention. In short, and yet again, “the alternative possibilities so nicely generated by the strategy of redescription of the content of moral responsibility lack robustness” (Fischer, p. 37).

16 In chapter 3, I provide an in-depth discussion of both this version of the flicker strategy and the ways in which it meets Fischer’s defeater (i.e., robustness objection).
§2.4.4 Analogue to Counterfactual Theories of Knowledge

Fischer attempts to improve his overall robustness objection with a clever analogy to epistemology, specifically to counterfactual theories of knowledge. Such theories traditionally claim that agents know a proposition $P$ only if they can track the truth of $P$ across multiple possible worlds. In other words, agents must be able to distinguish reliably between possible worlds at which $P$ is true and the possible worlds at which $P$ is false.

In what way do counterfactual theories of knowledge resemble PAP? PAP is consistent with the thesis that, in order for agents to be responsible for some action, they must be able to refrain from doing that action. So, the resemblance that Fischer proposes is that in both theses—PAP and counterfactual theories of knowledge—there is a necessary counterfactual condition. Suppose, for example, that Schmidt claims to know that there is a barn in front of him. Proponents of counterfactual theories of knowledge argue that Schmidt knows $P$ only if he can distinguish the actual situation (i.e., where there is a barn in front of him) from an alternative situation (i.e., where there is no barn in front of him). Further suppose, however, that Schmidt cannot distinguish the actual situation from the alternative situation. In other words, Schmidt believes there is a barn in front of him when in fact there is not. Schmidt has false beliefs in the alternative situation. According to Fischer, it would be odd for a proponent of such a theory to suggest that it is in virtue of the fact that Schmidt forms false beliefs about the barn in the alternative sequence that he has knowledge of the barn in the actual sequence. But, as

\[\text{17 This is the relevant implication of counterfactualist theories of knowledge.}\]
Fischer argues, this claim is directly analogous to the claim made by proponents of the flicker of freedom strategy in relation to Frankfurt-style cases. For, according to Fischer, they maintain that it is in virtue of an alternative sequence involving a coerced, non-free action that agents are responsible for their action in the actual sequence. In short, Fischer seems to think that the implausibility of counterfactual theories of knowledge infect the analogous flicker defense.

In light of the foregoing, Fischer believes that the flicker strategy is implausible. Moreover, it highlights the importance of Frankfurt-style cases. That is, they switch our focus to the actual sequence, and suggest that it is in virtue of the actual sequence, not the hypothetical alternative sequence(s), that agents are morally responsible. Thus, the task for proponents of PAP who appeal to the flicker of freedom strategy is to identify features of the alternative sequence that make agents responsible for their actions. Fischer is once again pressing the flicker theorist for a good reason why the mere existence of possibilities plays any explanatory role in agents’ moral responsibility.

§2.5 Conclusion

I began the chapter with a summary of the shift from traditional Frankfurt-style cases to modified Frankfurt-style cases. This shift is the result of the dilemma defense to PAP. Recall the initial problem with the example: either Jones’s twitch (or whatever the sign might be) is causally sufficient for Jones starting the fire, in which case the same thing that makes it impossible for Jones to start the fire is what causes Jones to start the fire (the deterministic horn), or Jones’s twitch is not causally sufficient for Jones starting the
fire, but merely accurately predicts a subsequent action, in which case it seems that Jones still has some alternatives available to him (the indeterministic horn). In light of this defense and the shift from traditional to modified Frankfurt-style cases, proponents of Frankfurt-style cases opened the door for the flicker of freedom strategy to PAP. There are at least four different versions of the flicker of freedom defense according to Fischer. According to him, none of these strategies met his robustness criteria. In other words, he presents a defeater that applies to all flicker defenses.

As a result of Fischer’s initial defeater, both flicker theorists and proponents of Frankfurt-style cases further examine the conditions for sufficiently robust actions. An adequate response to the robustness objection must demonstrate how it is that, even in Frankfurt-style cases, agents retain alternative possibilities that are sufficiently robust (i.e., there must be some way in which agents can actually do otherwise) in the sense in which Fischer requires. (Or, of course, one can simply reject robustness altogether.) I will discuss in the next chapter what constitutes a robust action.
Chapter 3

The Nature of Robust Alternatives

§3.1 Introduction

In Chapter 1, I presented a review of the debate surrounding PAP, both according to my interpretation and McKenna and Widerker (2006). There, I discussed the salient features of Frankfurt-style cases, the broader implications of the truth of PAP, and three different kinds of (modified) Frankfurt-style cases that have been offered since Frankfurt’s original case. One salient feature—namely, the fact that alternative possibilities are built into the examples—opens the door for a potential defense of PAP. This potential defense is the flicker of freedom defense strategy. Such a defense, in the very least, argues that there is a sense in which agents still could have done otherwise even in Frankfurt-style cases; thus, PAP intuitively remains unthreatened by this counterexample strategy.

In Chapter 2, I presented John Martin Fischer’s four versions of the flicker of freedom strategy, and his defeater to these strategies. Fischer presents an overall objection that can be applied to each version: the robustness objection. According to Fischer, the flicker of freedom\(^{18}\) strategy is ultimately unsuccessful because there is no robust sense in which agents in Frankfurt-style cases can actually do otherwise than what

\(^{18}\) The reader should take note that any instance of ‘flicker theorist’ (or any variation thereof) in this chapter is meant to refer specifically to Michael Robinson’s (2012 and 2013) fine-grained flicker strategy.
they in fact do. An adequate response to Fischer’s defeater must show that an agent’s alternative possibilities be sufficiently robust in the sense that they would have been able to do otherwise had they chosen to do so. It is therefore incumbent on the flicker theorist to offer a response to the robustness objection in order to preserve the truth of PAP.

Such a response to Fischer’s robustness objection is where I turn for the present chapter. While Fischer’s robustness objection claims that, in virtue of the existence of alternative possibilities, we are not morally responsible for our actions, others (Michael McKenna 2003, Michael Robinson 2012, and Seth Shabo 2013 and 2015) draw our attention to the significance of certain epistemic conditions as revisions to Fischer’s original objection. I contend that some of these revisions to robustness are too stringent of a requirement that flicker theorists need not satisfy.

I argue that flicker theorists need only respond to Fischer’s original defeater. The other requirements (McKenna 2003, Robinson 2012, Shabo 2013 and 2015) that emphasize the importance of satisfying certain epistemic conditions seem fruitless. Robustness requiring an epistemic condition seems wrongheaded. There are cases where we hold persons morally responsible for some action, even though they might not have knowledge of the significance of their action—namely, culpable ignorance. In this chapter, I will show how those (McKenna, Robinson, and Shabo) who advocate for an epistemic condition to robustness seem to miss the mark.
§3.2.1 Robustness Revisited

In both Chapters 1 and 2, I used the term “robustness” without much of an explanation of exactly what this means. Just as the debate surrounding PAP has developed, so too has this conception of robustness. This section will constitute a brief overview of the meaning of robustness.

In Chapter 2, I presented Fischer’s purported defeater of the flicker defense. As I discussed in Chapter 2, the alternative possibilities, \textit{qua} alternative possibilities, available to Jones in (both traditional and modified) Frankfurt-style cases are insufficiently robust in the sense that there is no way in which Jones could have actually done other than what he in fact did.

Since Fischer’s defeater, there have been a number of other Frankfurt defenders, as well as some proponents of PAP, (e.g., Robinson, Shabo, McKenna) who have cashed out this robustness requirement in an epistemic way. To be clear, recall that PAP says that people are morally responsible for what they have done only if they \textit{could} have done otherwise. According to Shabo (2015), there is room to explicate and defend “could” in an epistemic way. Though Shabo considers a slightly different angle to generate the relevant intuitions of power attributions as they pertain to the free will debate (that is, statements of the form “It’s up to Jones whether he does A instead of B…”), he contends that these statements involve a neglected epistemic condition (Shabo, 2015). In order to motivate this epistemic requirement of robustness, consider the following case that Shabo offers:
Turning Point: Chris is at an unexpected crossroads—or rather, a T-intersection. His mission is to deliver emergency supplies to a stricken village. Unfortunately, the person tasked with giving him directions failed to mention this particular juncture [sic]. Scanning the undifferentiated vegetation on all sides, Chris is forced to admit that he has no idea which way to turn. If he sets out in the wrong direction, it could be an hour or more before he realizes that he needs to backtrack. Unknown to Chris, the left fork leads to the village. (Shabo, 2015)

Importantly, there is nothing (internally or externally) influencing Chris to go a certain way; he has the ability to go right or left. I believe Shabo rightly contends that it is not up to Chris whether he turns the right way precisely because he is not in a position to promise someone that he will turn towards the village straightaway (Shabo, 2015).¹⁹

The main substantive theoretical claim Shabo motivates is that, from situations structurally similar to these, there are circumstances in which agents are not in a position to promise another that they will complete a certain course of actions, even if all their alternatives are enabling. According to Shabo, such circumstances rest on a “knowing that” and not a “knowing how” (Shabo, 2015). There is nothing Chris does not know how to do: he knows how to turn left, and how to turn right. What he lacks, however, is the propositional knowledge of which way will get him to the village. In a way it is not up to him that he turns towards the village. In another way, though, it is up to Chris that

¹⁹ Readers familiar with the luck objection to libertarians may notice that Shabo’s case is closely related to van Inwagen’s “Promising Argument” (2011). Van Inwagen’s aim is to demonstrate that agents who know that their courses of action, A and B, have roughly even objective probability are not in a position to promise A over B, or vice versa. This is allegedly problematic for libertarians as they are not in position to claim that they are able (in the libertarian sense) to perform causally undetermined actions. Shabo’s aim, on the other hand, is solely to account for the fact that Chris’s epistemic situation lacks the relevant ability (and not relevant abilities with respect to causally undetermined actions).
he turns towards the village. It is up to Chris that he turns towards the village, if the
statement refers strictly to his epistemic abilities. But Shabo contends that, in order for it
to be up to Chris that he turns towards the village, he must know that the left fork leads
towards the village.

He further concludes that, in addition to the essential information that Chris lacks,
agents have a “robust alternative possibility if and only if it is up to her whether that
possibility is realized” (Shabo, 2015). It is this feature, which I will argue shortly, that
turns out to be too stringent of an epistemically conditioned requirement to robustness.

§3.2.2 Robustness continued: McKenna and Robinson

McKenna (2003) contends that, in addition to the initial requirement put forth by Fischer,
a robust alternative must be such that “it would tell us something (different from what we
are told in the actual world) about the moral quality of the agent’s conduct were she to
have so acted in this alternative scenario” (p. 203). According to McKenna, an
alternative is robust only if the following two conditions are satisfied: (i) an alternative
course of action must be either not morally on a par or no worse than the action agents
actually perform—that is, it must be morally significant; and (ii) it must be reasonable to
expect that agents consider an alternative to what they actually do, given their agent-
relative perspective—that is, it must be deliberatively significant.

Robinson (2013) contends that Fischer and McKenna are right to claim that
robustness requires that agents have direct voluntary control over their action, but what’s
unclear is why robustness applies only to overt physical actions (p. 61). It seems that
Robinson’s point is that robustness need not apply exclusively to actions if the main reason for thinking so is because that is what the agents have direct voluntary control over. We seem to have just as much, if not more, control over decisions. So, I believe Robinson is right to question the nature of robustness and what exactly it requires. I will now discuss this question.

§3.3.1 Robinson’s Flicker of Freedom Strategy

Robinson (2012) maintains that proponents of PAP have available to them a flicker defense strategy to modified Frankfurt-style counterexamples. Recall that traditional Frankfurt-style counterexamples beg the question against incompatibilists, since the traditional structure of these examples includes a deterministic relation between agents’ free action and its causal antecedents. This is the line of defense Robinson takes, and this is also the one I endorse in this chapter. Proponents of such cases have reformulated them in such a way so as not to be guilty of begging the question against incompatibilists.

The attempt to reformulate Frankfurt-style cases in a non-question-begging way, however, allows for a revival of the flicker of freedom defense strategy of PAP. Even though Chapter 2 made distinctions among four versions of the flicker strategy, this does not mean that the shift from traditional to modified Frankfurt-style counterexamples allows for a revival of all flicker strategies. Although he does not explicitly comment on the issue, I believe Robinson and I agree that the shift in Frankfurt-style cases allows for

---

20 The objection to traditional Frankfurt-style counterexamples does not end there. This is merely one horn of the dilemma. The other horn, then, is that, if it is not a deterministic relationship between Jones’s free action and its causal antecedents (i.e., if it is not the first horn of the dilemma), then it must be merely probable that Jones does in fact kill Smith. But, if it is merely probable that Jones will kill Smith, then it Jones could still have done otherwise than kill Smith and, consequently, PAP remains unthreatened.
the success of one particular version—namely, the fine-grained version. Integral to this strategy are two questions. First, what precisely is entailed by Robinson’s proposed flicker defense strategy? And, second, given its notorious ambiguity, what is PAP really a thesis about? In other words, what is the relationship between moral responsibility and alternative possibilities?

Robinson argues, among other things, that PAP, in the very least, is a thesis about one necessary condition for agents’ moral responsibility. In particular, this necessary condition is that agents must be able to otherwise. Other than this, however, it is fair to say that PAP is a vague principle. Does PAP apply exclusively to mental modes (i.e., decisions, beliefs, etc.) or overt physical actions (i.e., actually doing some action)? Does it apply to consequences of actions or merely to actions themselves? In order to bring to the fore what exactly this vagueness is, consider the following example. Suppose a terrorist organization’s bank is located in an area that includes thousands of innocent people. In order to stop the ongoing terrorist attacks, we must cut off their financial supply—i.e., we must blow up the bank. The bank’s destruction, however, will come at the cost of the lives of all the (innocent) people in the surrounding area. Further suppose that the burden is on me to push the button that will cause a missile to strike the bank. I decide on my own to push the button. So, I push the button, the bank is destroyed, innocent people die, and I caused it all. But, among other things, the pertinent question to

---

21 The reader should note that even this relationship is subject to dispute. For instance, Garry Young (2016) argues that, rather than PAP being understood as a necessary condition for agents’ moral responsibility, it is a sufficient condition. He proposes that, if we understood PAP alternative possibilities as being a sufficient condition for moral responsibility, we have a principle that is immune to Frankfurt-style cases. Although this dialectical move may very well be promising, I am working under the assumption that alternative possibilities is a necessary condition for moral responsibility.
ask: in what way am I responsible for each of these occurrences? Surely I am not morally responsible for pushing the button in the same way that I am morally responsible for the bank’s destruction or the death of all the innocent people. It seems that I am only derivatively morally responsible for the death of the innocent people and the bank’s destruction, and that I am basically (i.e., directly or non-derivatively) morally responsible for deciding to push the button. Again, this distinction is not meant to absolve my culpability for the situation in general, but rather to get clear on the sense in which I am morally responsible for the event (and each subsequent event) to my pressing the button.

The difference between my derivative and basic responsibility lies in the fact that my moral responsibility for the bank’s destruction and the death of the innocent people derives from my pushing the button. Call this example “Adam’s Bank.”

It is still difficult to see the sense in which Adam’s Bank pertains to PAP such that the relationship between moral responsibility and alternative possibilities is clear. I contend that Adam’s Bank (and any example structurally similar to it), however, provides the first step in understanding what PAP is a thesis about. The distinction between derivative and basic responsibility is extended to mental modes (e.g., decisions) and their subsequent overt physical actions. For instance, the fact that I freely\textsuperscript{22} raise my arm is due to my decision to raise my arm. In this sense (which is structurally similar to Adam’s Bank), then, my responsibility for raising my arm derives from my responsibility for deciding to raise my arm: I am basically responsible for deciding to raise my arm.

---

\textsuperscript{22} By “freely” I mean nothing more (nor less) than simply intending to do a certain thing, uninfluenced by any external factors.
It is in this respect that there is a connection between PAP and traditional Frankfurt-style cases. To put it in the context of Post Office (see §2.2.1), according to how things actually happened, Jones’s moral responsibility for starting the fire derives from his decision to set fire to the post office on his own. In other words, Jones is basically morally responsible for deciding on his own to burn the place down, and causal consequents are the basis for derivative responsibility (that is, setting the post office on fire is due to his decision). Despite Frankfurt’s overall stance on the issue, it seems even he agrees with this distinction (Robinson 2012, p. 187). As Frankfurt writes, “[even though Jones] has no alternative but to do what Black wants him to do … If he does it on his own … his moral responsibility for doing it is not affected by the fact that Black was lurking in the background with sinister intent, since this intent never comes into play” (1969, p. 836). Even though Frankfurt is not explicitly arguing for this distinction between derivative and basic moral responsibility, the implication is that “had Jones not acted on his own, he would not be morally responsible for what he did” (Robinson 2012, p. 187). What it is to be basically morally responsible for an action can be characterized in the following way:

PAPb: People are basically morally responsible for what they have done only if they could have done otherwise (Robinson 2012, p. 188).

It is in this sense, then, that Frankfurt seems to agree that PAP is, first and foremost, a thesis about basic moral responsibility. If Jones had not decided on his own to set fire to the post office, though he would still have done it, he would not be morally responsible (in any sense) for his actions.
It is easier to see the implications of this distinction as it pertains to Post Office, whereas it is more difficult to make sense of the way in which this distinction can aid defenders of PAP in Modified Post Office (see §2.2.1). In Post Office, Jones will either set fire to the post office of his own volition or he will set fire to the post office due to Black’s intervention. Jones will set fire to the post office either way; it’s simply a matter of how he goes about doing so. According to how the series of events actually unfolds, Jones sets fire to the post office on his own, and it is for the decision that he is basically morally responsible. Now, because of his ability to set the post office ablaze either of his own accord or Black’s, Jones indeed has some degree of control over his actions (even though he has no knowledge of Black’s presence). Jones can control the way in which he sets fire to the post office. Thus, even in Post Office, Jones possesses alternative possibilities in the sense that he can still do otherwise than set the post office on fire on his own and, consequently, it is in this sense that traditional Frankfurt-style cases do not falsify PAP. Thus, it seems that understanding PAP in terms of basic moral responsibility brings to the fore the extent of Jones’s control in Post office. For, what Jones does have control over is whether or not he burns the place down on his own or due to Black.

Regarding Modified Post Office, Robinson rightly claims that “Jones has the same amount of control over whether he will set fire to the post office as he does over whether he will decide to set fire to the post office—namely, none” (Robinson 2012, p. 187). However, simply because Jones does not have control over his action or his decision does not automatically preclude Jones from being basically morally responsible—that much was shown by Frankfurt’s original case. In Modified Post
Office, Robinson shows that Jones is basically morally responsible for deciding on his own to set the post office ablaze only if he could have done otherwise than decide on his own to set fire to the post office (Robinson 2012, p. 189). It follows, then, that “Modified Post Office will be successful in falsifying PAP [as understood by “basic” moral responsibility] only to the extent that it is true that Jones is basically morally responsible for deciding on his own to set fire to the post office and yet false that Jones could have done otherwise than decide on his own to set the post office on fire” (Robinson 2012, p. 189).

§3.3.2 Robinson on Robust Flickers of Freedom

So far, it seems clear that any successful flicker strategy must do (at least) two things. First, (i) it must address the robustness objection as initially presented by Fischer. Second, (ii) it must demonstrate how these flickers are in fact robust in the sense required by Shabo. Regarding the first, it is clear that both flicker theorists and those who reject PAP accept something like the following principle, which is a closely related to PAP. Call it the Principle of Robust Alternative Possibilities (PRAP):

\[ \text{PRAP: Individuals are basically morally responsible for what they have done only if a robust alternative to it was available to them.} \]

---

23 It seems to me that an adequate response to (i) would render (ii) obsolete. The reason doing (i) might obviate the objection is if one restricts the scope of their project to merely responding to Fischer’s defeater. My aim, however, is not only to defeat Fischer's defeater, but it is also to show Shabo’s stringent epistemic requirement begs the question against flicker theorists.
Robinson offers the following argument to demonstrate that modified Frankfurt-style cases, specifically indeterministic ones (e.g., Modified Post Office), are theoretically incapable of satisfying the counterexample requirements for PRAP:

1) PRAP is a thesis about basic moral responsibility. (Robinson, p. 63)

2) A counterexample to PRAP must be a case in which individual(s) had no robust alternative to doing that for which they are basically morally responsible. (Robinson, p. 63)

3) Agents in Frankfurt-style cases are basically morally responsible for deciding on their own to act as they do (not for so deciding *simpliciter*). (Robinson, p. 63)

4) A counterexample to PRAP must be a case in which a person had no robust alternative to deciding on his own to act as he did. (from 2 and 3) (Robinson, p. 64)

5) An alternative is robust if and only if it is (i) morally significant, (ii) deliberatively significant, (iii) understood to be responsibility-precluding, and (iv) within an agent’s voluntary control. (Robinson, p. 65)

6) Omitting to decide on one’s own to do that which one *knows* to be wrong is a morally significant alternative to so deciding on one’s own (in all Frankfurt-style cases). (Robinson, p. 65; italics my own)

7) Omitting to decide on one’s own to do that which one *knows* to be wrong is a deliberatively significant alternative to so deciding on one’s own (in all indeterministic Frankfurt-style cases, e.g., Modified Post Office). (Robinson, p. 66; italics my own)
8) Omitting to decide on one’s own to perform an action is understood to preclude one from being (basically) morally responsible for so deciding (in all Frankfurt-style cases). (Robinson, p. 66)

9) Voluntarily omitting to decide on one’s own to do that which one knows to be wrong is within one’s voluntary control whenever one voluntarily omits to do so. (Robinson, p. 66)

10) Voluntarily omitting to decide on one’s own to do that which one knows to be wrong is a robust alternative to so deciding on one’s own (in all indeterministic Frankfurt-style cases). (from 5–9) (Robinson, p. 67; italics my own)

11) It is an essential feature of indeterministic Frankfurt-style cases that it is possible for agents in these cases voluntarily to omit to decide on their own to do that which they know to be wrong. (Robinson, p. 67; italics my own)

12) It is an essential feature of indeterministic Frankfurt-style cases that agents in these cases have robust alternatives to deciding on their own to act as they do. (from 10 and 11) (Robinson, p. 68)

13) Therefore, indeterministic Frankfurt-style cases are theoretically incapable of serving as genuine counterexamples to PRAP. (from 4 and 12) (Robinson, p. 68)

From this argument, Robinson further concludes:

So long as it remains open to agents in these cases to knowingly and voluntarily omit to decide on their own to do that which they know to be wrong, they will continue to possess robust alternative possibilities. And since there is no way to eliminate the possibility of agents knowingly and voluntarily omitting to decide on
their own to act as they do without undermining the judgment that they are morally responsible for their actions, agents in these cases will continue to possess robust alternative possibilities as long as they are morally responsible for what they do. Thus, either agents are not morally responsible for what they do in these cases or they continue to possess robust alternative possibilities. Either way, these cases will fail to falsify PRAP. (Robinson, pp. 68-69; italics my own)

Note that Robinson’s argument here involves *omissions* and not *actions*. Despite the fact that it does not concern actions, I believe Robinson’s main point still applies because PRAP (and so PAP) is a thesis about both actions and omissions. Of course, this is not to say that either principle is solely about actions and omissions. I still contend that, insofar as PAP (and PRAP) is a thesis about actions, it must also be a thesis about mental actions (i.e., decisions), as expressed by how individuals can be basically morally responsible for their actions. In sum, I believe Robinson is correct to conclude that indeterministic Frankfurt-style counterexamples are incapable of falsifying PAP.

§3.4.1 McKenna (and Robinson) on Robust Alternatives

In the previous section (§3.3.2) and §3.2.2, I distinguished between Fischer’s robustness objection and others’ (McKenna 2003 and Robinson 2012) emphasis on the importance of satisfying an epistemic condition. Fischer requires that robust alternatives be within agents’ voluntary control, and others include that certain epistemic conditions be met.

---

24 More specifically, it seems Robinson implicitly means to say that PRAP is a thesis about actions, *intentional* omissions (i.e., refraining). What seems clear is that PRAP is not a thesis about *unintentional* omissions (or actions).
Fischer and McKenna seem to presuppose that a robust alternative consists in action. Robinson agrees with this but adds a clarification: he contends that mental actions serve as the basis for attributions of morally responsible agency. Again, the reason for this is that, in order for an alternative to be robust, it must be within the agent’s direct voluntary control, and (assuming a situation with no manipulation that is unknown to the agent) decisions are the things that agent’s have direct voluntary control over.

I believe Robinson is right to contend that decisions are within an agent’s direct voluntary control. But it does not seem that he also needs to accept McKenna’s proposed epistemic conditions. To be clear, the success of Robinson’s argument—namely, that indeterministic Frankfurt-style cases are incapable of falsifying PAP—depends on his characterization of robustness including an epistemic condition. Robinson’s argument, however, is not successful. Premise (5) states that “An alternative is robust if and only if it is (i) morally significant, (ii) deliberatively significant, (iii) understood to be responsibility-precluding, and (iv) within an agent’s voluntary control” (Robinson, p. 65). But (iv) discounts Modified Post Office: Jones does not possess voluntary control over his alternatives—the indeterministic stipulation precludes this. In short, condition (iv) of premise (5) cannot be satisfied in Modified Post Office, and Robinson has to bite the bullet and accept that his argument does not defend the truth of PAP (with regards to Modified Post Office).

To put the worry in a slightly different way, consider premise (7), which states that “Omitting to decide on one’s own to do that which one knows to be wrong is a deliberatively significant alternative to so deciding on one’s own (in all indeterministic
Frankfurt-style cases)” (Robinson, p. 66). Here, what is deliberatively significant is an empirical claim and depends on whether an agent actually considers alternatives as such. In Modified Post Office, Jones is unable to deliberate in such a manner because he does not know the alternatives he has available. Moreover, premise (7) echoes McKenna’s robustness characterization. McKenna has two conditions, one of which is that an alternative is robust only if it is deliberatively significant. By this, McKenna means that it must be reasonable to expect that agents consider an alternative to what they actually do, given their agent-relative perspective. In other words, consideration of the alternatives to what an agent actually does is essential to deliberative significance.

This condition, however, seems too strong. Agents who consider the alternatives to what counterfactually could have occurred is just as significant to what they actually do. That is, McKenna’s condition can be counterfactual rather than actual and still be significant. Consider the following example. Michelle is on her way home from work. She is hungry, tired, and irritated. It’s a normal workday for Michelle. Due to the lack of food at her home, she makes a stop at Chipotle for dinner. She notices that the person in front of her is not paying attention to the line. Michelle could easily steal this person’s spot in line without any consequences (the person would not notice). She decides to take the person’s spot in line. According to McKenna, Michelle’s alternative is robust only if Michelle considers an alternative to what she actually does, and thus deliberative significance is satisfied. (Of course, this is in addition to the satisfaction of McKenna’s morally significant condition. Suppose, for the sake of the example, that this is also satisfied.) I contend, however, that Michelle need not consider an alternative to what she
actually does. She merely needs to consider the consequences of both her alternatives—namely, taking the person’s spot in front of her or not. Thus, rendering McKenna’s deliberatively significant condition as counterfactual, rather than actual, is still significant.

§3.4.2 Too Stringent of an Epistemic Condition

In §3.2.1, I presented Shabo’s characterization of robustness. Recall that he posits that “someone has a robust alternative if and only if its realization is up to her and that this type of power attribution is epistemically conditioned” (Shabo, 2015).

Moreover, according to Shabo, recall that Turning Point demonstrates that agents must possess all relevant propositional knowledge if they are properly said to attain their goal. In order for Chris’s alternatives to be robust, he must know that turning left will lead to the village. The example stipulates, however, that not only does Chris not know which way will get him to the village (or that there is an intersection, for that matter), there does not seem to be any way in which Chris can know this. Indeed, if Chris simply had a map, he would be able to reach the village. The stipulation precludes Chris from having (or even being able to acquire) the pertinent propositional content needed to get to the village. According to Shabo, Chris fails to satisfy robustness as the right half of the biconditional is false. In other words, it is not up to Chris whether he turns the right way, so it is not true that Chris has a robust alternative. But, if Chris had possessed the relevant knowledge, then it would have been up to him. There is nothing precluding this counterfactual other than Shabo’s stipulation; it is not as if Chris does not possess the
ability to retain such knowledge. If Chris had lacked the *ability* to retain such
knowledge, then it seems he would lack alternative possibilities altogether. If there were
no alternative possibilities available to Chris, there would not be any analogous features
to Frankfurt-style cases. It does not seem that such a case would shed any light on the
debate surrounding PAP. Recall that in §1.3.3, I claimed that there are certain
alternatives built into Frankfurt-style cases. Moreover, if these examples include an
individual who did not have any alternatives available, then the example begs the
question against incompatibilists (§1.3.5). To be sure, Turning Point is not intended as a
counterexample to PAP, but rather as a thought-experiment intended to motivate Shabo’s
intuition for what it is for an agent to have a robust alternative. So, the sense in which it
fails to shed light on the debate surrounding PAP is that, while Frankfurt-style cases have
alternatives built into the examples, and there is no stipulation precluding an agents
ability to act (or decision to act), Turning Point stipulates a stringent epistemic condition
that precludes an agents ability to do certain things—namely, acquire the relevant
propositional content. It seems, then, that Shabo’s claim that robustness is epistemically
conditioned is insufficiently motivated, since there is no adequate support for Turning
Point’s crucial stipulation. It is in this sense, I believe, that Shabo is demanding too much
from the flicker theorists.

I will put my concern in a slightly different way. Shabo further claims that he has
“offered a negative characterization to the effect that the agent does not meet the
condition if she has no conscious awareness of the relevant facts about her alternatives or
of how to discover these facts” (2015). Regardless of whether Shabo’s aim is to offer a
negative or positive characterization, I contend that this condition is too stringent. This requirement is too stringent in the sense that it need not be expected that agents should have always understood the significance of their actions. What seems sufficient is only that agents should have understood their action(s) and the ways to complete them, so long as they were able to understand and the information was available to them (see Timpe 2013, pp. 100-110). Such cases include instances of culpable ignorance where the agent does not actually understand.

§3.5 Conclusion

Due to the conditional nature of PAP, defenders may either argue that agents are not morally responsible in Frankfurt-style cases or, even in Frankfurt-style cases, agents still retain some alternative possibilities. This thesis takes the latter line of defense. I argue that agents in Frankfurt-style cases still possess some alternatives. Coined by John Martin Fischer, this is the flicker defense strategy of PAP. Fischer initially accused the flicker theorists’ line of defense as being unsuccessful, since the alternatives under consideration are insufficiently robust: the mere existence of alternative possibilities being available to agents is not enough to ground attributions of moral responsibility. I argue that certain attempts to meet Fischer’s defeater are successful, in particular the fine-grained flicker strategy version developed by Michael Robinson (2012 and 2013). Other attempts at revising the robustness requirement, in particular from Seth Shabo (2013 and 2015), are unsuccessful in undermining the flicker theorists’ strategy.
Conclusion

In his article, “Alternate Possibilities and Moral Responsibility,” Harry Frankfurt offered a class of putative counterexamples to the Principle of Alternative Possibilities (PAP) which states that persons are morally responsible for what they have done only if they could have done otherwise. In this thesis, I narrowed in on some of the details in the current debate.

In Chapter 1, I discussed some of the motivations for believing in PAP, as well as the broader philosophical implications. Then, I explained Frankfurt’s challenge to PAP and then I explained the salient features of Frankfurt-style cases. I then discussed the dilemma defense: due to the conditional nature of PAP, a defense must show that, in the proposed counterexample, either individuals are not morally responsible, or they have alternatives still available. Lastly, I discussed three different kinds of Frankfurt-style cases that emerged in response to the dilemma defense.

In Chapter 2, I focused on another line of defense for PAP: the flicker defense. Flicker theorists’ strategy, among other things, is to show that the existence of these prior signs in Frankfurt-style cases is enough to ground attributions of moral responsibility. John Martin Fischer developed four versions of the flicker defense, and then proposed the robustness objection to the four versions. As a result of Fischer’s defeater, flicker theorists must show how an agent’s alternatives can be sufficiently robust.
Chapter 3 argued that, by deploying the notion of basic responsibility, flicker theorists can adequately respond to Fischer’s robustness objection. I showed how arguments put forth by McKenna, Robinson, and Shabo for an epistemic robustness condition missed the mark. In doing so, I argued that the further epistemic robustness conditions were unrealistic and too stringent.

I began this thesis with the ambition of offering a defense of PAP without endorsing any larger philosophical theories within the free-will debate. In a sense, I failed to do both of those things. I did not offer a defense of PAP. Rather, I defeated a defeater. If it happens to turn out that PAP is true, all the better for proponents of PAP. It is also unclear that its truth is because of my defeater. In another sense, though, I achieved both of these things: PAP still is true in regards to certain Frankfurt-style cases, and I did not claim enough about a notion of free will to be considered either a compatibilist or incompatibilist.
Bibliography


Pereboom, Derk. 2003. “Source Incompatibilism and Alternative Possibilities,” in

Widerker and McKenna (2003), 185-199.


*Philosophical Explorations* 12: 109-118.


*Philosophical Studies* 160: 429-446.


