

compos mentis

Moral Responsibility in the Age of Free Will Skepticism: A Defence of Frankfurtian-Compatibilism

Owen Crocker

University of Victoria

ABSTRACT

Free will skepticism is radical in its core claim that free will is illusory. Criminal law, however, appears to presuppose that persons are free and hence, morally responsible for their actions. So, if free will skepticism is true, our current practices that hold people to account for their wrongs appears unjustified—even immoral. This paper will challenge the free will skeptic's core claim that free will does not exist and defend current practices of moral responsibility by offering (and defending) a Frankfurtian-compatible approach to the topics of free will and determinism.

KEYWORDS

Free Will, Moral Responsibility, Crime, Free Will Skepticism, Criminal Law, Compatibilism, Determinism, Libertarianism, Legal Personhood, Autonomy, Rationality

INTRODUCTION

Crime, free will, and moral responsibility are socially and philosophically loaded terms, which often convey different meanings to people. The criminal justice system utilizes and refers to many of these terms to justify its punishment for those found guilty of criminal activity.

The practice of holding citizens morally responsible for their behaviour is a common social phenomenon, that most do not seriously question. However, according to free will skeptics (those who argue that free will does not exist), social practices that hold persons morally responsible are unjustified and in need of reform. The goal of this paper is to demonstrate that the legal practice of holding people morally responsible can be justified using a Frankfurtian-compatibilist approach.

The paper will be divided into three parts. The first part will outline three presuppositions of the criminal law. The second part of the paper will provide the reader a short summary on free will skepticism. The third part of the paper will present an alternative approach to free will skepticism. That approach being, Frankfurtian-compatibilism. In this section, Frankfurtian-compatibilism will be described as well as defended against two possible objections.

PART ONE: CRIMINAL LAW'S PRESUPPOSITIONS: LEGAL PERSONHOOD, AUTONOMY AND RATIONALITY

In this portion of the paper, three presuppositions of the criminal justice system¹ will be discussed: legal personhood, autonomy, and rationality. It seems to me that all three of these concepts play an integral role in the criminal justice system's justification in apportioning moral responsibility.

Regarding legal personhood, Naffine writes, "...the law of persons is not a discrete field of study in the common law world, such as torts, or contract or criminal law, but is a *pervasive underlying concept throughout the different branches of law*" (2009, 15) [Emphasis Added]. Criminal law, like other legal disciplines, all

1. It should be noted that the author is writing from a Canadian perspective. In this regard, references to case law or criminal practices are referring to Canadian law and practices. Many of the themes and elements will be recognizable in other criminal justice systems, but there will likely be some distinguishing factors between various countries methods of justice.

necessarily utilize what is termed, legal personhood. Without legal personhood, there would be no persons to convict of crimes, or for that matter, persons who could facilitate the court processes of convicting and acquitting those accused of crimes. The criminal law's person is comprised of many factors in assessing one's level of moral responsibility for their behaviour. For example, a 5-year-old, in the eyes of the law, is not the proper subject of punishment for breaking and entering, whereas a 27-year-old is. In addition to age, the criminal law also presupposes that the legal person is of sound mind, demonstrating the contrary would excuse a person by means of the *not criminally responsible due to mental disorder* defence. However, if such a defence is not raised or not considered, it is assumed that the accused is "normal" regarding their cognitive abilities and decision-making processes. The law of persons, as Naffine notes, is an *underlying* concept of law (2009, 1). Legal personhood can be seen as the base conceptual unit (Naffine, 2009, 1), which grounds the other faculties and powers required by the court, including autonomy and rationality.

Autonomy, meaning self-rule, is another underlying assumption in how persons are viewed and judged by the court. The ability to self-govern and control one's course of action in life is an incredibly powerful and entrenched thought in liberal societies (Nedelsky, 2011/2012, 121). As Ferguson writes, the criminal law presupposes "that human beings are rational and *autonomous*. We have the capacity to reason right from wrong, and the capacity to *choose right or wrong*. These assumptions may be incorrect, but they are, and are likely to remain, the theoretical basis of our criminal law" (1989, 140). When considering autonomy, the court will take into consideration the circumstances and factors surrounding a person's actions to determine if they were acting in a truly autonomous manner.

As Ferguson noted above, the criminal law assumes that persons are *rational* in that we possess the ability to *reason* right from wrong (1989, 140). Those deemed to lack rationality, such as children and those who suffer from mental disorders, are generally excused from being held liable, or are at least held to a lesser degree of responsibility. In the philosophical literature, the ability to reason right from wrong is often referred to as *reason-responsiveness*. Duff describes reason-responsiveness as,

The capacities on which responsibility depends are best understood as a matter of reason-responsiveness: a responsible agent is one who is capable of recognising and responding to

compos mentis

the reasons that bear on his situation. A responsible agent is 'responsible'² to reasons: which means not that he is responsible only when and insofar as he is actually responsive to reasons (since we can be responsible for our very failures to respond to relevant reasons), but that he is responsible insofar as he is capable of responding appropriately to relevant reasons. (2007;2009, 39)

The ability to reason right from wrong and to weigh the relevant factors before one chooses³ is a morally salient element that the courts will require in determining one's culpability.

At this point, it should be clear that the criminal justice system views *certain* persons as the appropriate subjects to be held morally responsible. As noted in this section, not all persons are appropriate candidates, as the elements of legal personhood, autonomy and rationality must be present. When these features are shown to be absent, the court will either not hold the person morally responsible or lower the level of their responsibility in the matter. However, all these distinctions and practices that are made in the criminal justice system take place because there is the general belief that persons are morally responsible for their actions (or that persons are at least in principle capable of being held morally responsible for their actions). However, not all persons share this belief. In the following section, free skepticism will be introduced.

PART TWO: FREE WILL SKEPTICISM

Free will skepticism, according to one of its lead proponents, Gregg Caruso, can be described as a viewpoint that "takes seriously the possibility that human beings are never morally responsible in the basic desert sense" (2021, 8). In other words, persons lack the necessary control to make themselves the proper subjects to be truly deserving of blame or praise for their actions. As will be argued below, the legal concepts of autonomy and rationality appear to make two assumptions.

-
2. This term is used by Duff to mean "able to respond" to reasons, hence the different spelling.
 3. Or the lack thereof of demonstrating sufficient rationality which the Court would consider negligence, wilful blindness, recklessness etc. For further discussion, see Sarch, A. (2019). *Criminally ignorant: Why the law pretends we know what we don't*. Oxford University Press at page 29.

First, that the agent in question could have acted differently than they did. Second, that agents are the source of their actions in a way that allows them to be morally responsible. However, if causal determinism is true, according to many free will skeptics,⁴ we have strong reasons to doubt both assumptions.

Philosophical viewpoints don't arise out of thin air. Rather, they are the culmination of many philosophical, social, and scientific theories. In the case of free will skepticism, there are essentially "two routes" that may lead someone to accept the free will skeptic's conclusion. The first of these routes relies on scientific findings. Specifically, the research conducted by Benjamin Libet was very influential to the free will skeptic's cause. In the 1980's, Libet ostensibly demonstrated the absence of free will. As Caruso and Pereboom write,

The pioneering work...by Benjamin Libet and his colleagues... investigated the timing of brain processes and compared them to the timing of conscious will in relation to self-initiated voluntary acts and found that the conscious intention to move...came 200 milliseconds before the motor act, but 350–400 milliseconds after the readiness potential (RP)—a ramplike buildup of electrical activity that occurs in the brain and precedes actual movement. (2018, 195)

Stated differently, our conscious will to move our body in a specific way or speak a specific phrase was happening *after* the electrical activity began to perform that specific action. Libet-type experiments are intended to demonstrate that when a person makes a choice to act in a certain way, that choice is not a result of their personal deliberation, but rather stems from an unknown prior physical cause. If this is true, then such findings call into question the assumption that persons are the true source of their actions. However, such scientific findings are questionable for various reasons (Caruso and Flanagan, 2018, 196-197).⁵ In addition to scientific

-
4. Many free will skeptics argue that free will does not exist regardless of whether it turns out determinism or indeterminism is true. However, for the purposes of this paper, determinism will be assumed.
 5. For example, one can concede that ramp-like electric activity builds up before a conscious decision is made but does not need to agree that such electric activity determines how one acts. It is possible that during or after the electric buildup, one makes the decision regarding how they will act. In this sense, the electric activity is simply a precursor to being able to act, not a determinant in how one acts.

compos mentis

reasoning, metaphysical speculations have also played an important role in the reasoning of free will skeptics.

The second route that may lead one to accept free will skepticism is premised on the idea of determinism. Such philosophical speculations can be traced all the way back to the Stoics in the 3rd century BCE (McKenna and Pereboom, 2014, 263). Determinism is the idea that every event that occurs can be fully explained by previous causes, whether those causes be social or biological. As Dennett writes, "If determinism is true, then our every deed and decision is the inexorable outcome, it seems, of the sum of physical forces acting at the moment, which in turn is the inexorable outcome of the forces acting an instant before, and so on, to the beginning of time" (2015, 5). Essentially, according to determinists, our choices and decisions are not ultimately *rooted* in our agentic powers, but rather are necessitated by prior causal forces that precede our every action and thought. If this deterministic picture is correct, it is difficult to see how a person could have acted differently than how they did in any prior circumstance. If everything is determined beforehand, how can persons be held responsible if they couldn't have acted otherwise than how they did? In comparison to the scientific findings of Libet which focused on physical processes in the brain, deterministic theories tend to focus on the metaphysical implications of a materialistic universe on free will and moral responsibility. Therefore, assuming our actions are determined, the criminal law's practice of holding people morally responsible for their actions becomes increasingly questionable.

PART THREE

Frankfurtian-Compatibilism

In contrast to free will skepticism, a viewpoint which holds the incompatibility of determinism and free will (as well as moral responsibility), compatibilists hold that determinism and free will are *compatible* with one another. There are many variations of compatibilism—all offering different accounts of what constitutes free will and moral responsibility. As stated at the beginning of this paper, the Frankfurtian-compatibilist account will be offered as an approach that can justify the current legal practices of holding persons morally responsible for the crimes they commit.

Frankfurtian-compatibilism is premised on what are called, first and second-order desires. First-order desires are those desires which have objects as their goal. For example, eating a delicious meal or getting a good night sleep. On the other hand, second-order desires have other desires as their goal. In other words, they are “desires about desires” (McKenna and Coates, 2021). For instance, one might think to herself, that it would be great to have the motivation to study for a big test. In this sense, the subject is desiring the desire to study. The ability to be able to reason between our first and second-order desires as well as supplant previous desires with new ones is the foundation of Frankfurt’s conception of free will. Compared to other animals, humans are unique in their ability to possess second-order desires which give people reasons not to carry out first-order desires. Frankfurt writes, “the statement that a person enjoys freedom of the will means (also roughly) that he is free to want what he wants to want. More precisely, it means that he is free to will what he wants to will, or to have the will he wants” (1971, 16). It is an important aspect of Frankfurt’s theory that persons possess the power to change or have control over their desires. It should be noted that Frankfurt acknowledges that people’s desires and how they choose to act considering those desires are “far more complicated” than the first and second-order paradigm (Frankfurt, 1971, 16). There is potentially no limit to how many higher order desires one might have, which can quickly complicate one’s reasoning processes. However, Frankfurt’s outline of free will coupled with the control of adding or replacing desires seems to be a sufficient grounding for how moral responsibility is attributed to people. Frankfurtian-compatibilism will be further strengthened by responding to two common objections. The first of those being, the principle of alternate possibilities.

The Principle of Alternate Possibilities

The first argument that will be considered is titled, *the principle of alternate possibilities*. When considering questions of free will and moral responsibility, the openness of the future and an agent’s ability to choose their actions is of great importance. As stated in section two, determinism is the idea that “our every deed and decision is the inexorable outcome...of the sum of physical forces”. So, if one views an accused’s decision to commit a crime as the inexorable outcome of physical forces, then intuitions of their guilt and responsibility quickly disappear. This reasoning can be formalized into the following principle:

compos mentis

PAP: A person is morally responsible for what she has done only if she could have done otherwise (Kane, 2005, 283).

As Haji notes, those who hold to PAP, “insist that alternative possibilities are required for the active control an agent must exercise in performing an action for the action to be free. They claim that determinism undermines free action or responsibility because it undermines active control by eliminating alternatives” (2012, 190). However, PAP along with its implications, is not without its dissenters.

Possibilities versus Reasons

In response to the principle of alternate possibilities, Frankfurt published a paper titled, *Alternate Possibilities and Moral Responsibility* (1969), in which he attempts to demonstrate why the principle of alternate possibilities is mistaken. In the paper, Frankfurt utilizes an *intuition pump*,⁶ which is intended to make us question whether the ability to do otherwise is necessary for persons to be morally responsible. Frankfurt writes the following:

Suppose someone - Black, let us say - wants Jones to perform a certain action. So, he waits until Jones is about to make up his mind what to do, and he 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. What steps will Black take, if he believes he must take steps, in order to ensure that Jones decides and acts as he wishes? Let Black give Jones a potion, or put him under hypnosis, and in some such way as these generate in Jones an irresistible inner compulsion to perform the act Black wants performed and to avoid others.

Now suppose that Black never has to show his hand because Jones, for reasons of his own, decides to perform and does perform the very action Black wants him to perform. In that case, it seems clear, Jones will bear precisely the same moral responsibility for what he does as he would have borne if Black had not been ready to

6. For more on intuition pumps, see Dennett, D. C. (2013). *Intuition Pumps and Other Tools for Thinking*. W.W. Norton & Co.

take steps to ensure that he do it. It would be quite unreasonable to excuse Jones for his action, or to withhold the praise to which it would normally entitle him, on the basis of the fact that he could not have done otherwise. Indeed, everything happened just as it would have happened without Black's presence in the situation and without his readiness to intrude into it. Jones has no alternative but to do what Black wants him to do. If he does it on his own, however, 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, 835-836).

As one can see in Frankfurt's example, Jones appears to be acting in a *voluntary* manner. His decision to act is in line with his own desires and without outside interference. In this sense, Jones' act is free and is intuitively deserving of praise and blame. There appears to be little to no reason to conclude that Jones' behaviour is involuntary, even though he was unable to do otherwise. The Jones example is supposed to demonstrate that when determining if one should be held responsible for their actions, the salient moral criteria is not whether they could've acted otherwise, but instead what their *basis* and *reasons* were for acting in the manner they did. Such moral reasoning is made clear in McKenna's following principle:

L-Reply: A persons' moral responsibility concerns what she does do and her basis for doing it, not what else she could have done (McKenna and Pereboom, 2014, 104).

The L-Reply, like Frankfurtian-compatibilism, looks to the reasons (or basis) for why one acts as they do. It seems to me that this approach is preferable to PAP, as PAP will unjustifiably rule out many cases where one freely choose to act in a certain way. For example, if Jones' act was criminal in nature, it appears to satisfy all the relevant criminal law considerations. He acted voluntarily⁷ (*actus reus*), intentionally (*mens rea*) and had no apparent defense to explain his behaviour. Therefore, Jones appears to have made a free choice, that he should be held responsible for. In light of these reasons, if the courts were to look at the *reasons*

7. Voluntarily in this sentence means physical voluntariness, not moral voluntariness.

and *basis* for one's actions, they would find Jones morally responsible. In contrast, however, if the courts applied PAP, Jones would not be responsible, since he couldn't have done otherwise. Intuitions may differ at this point, but it seems far from clear that Jones is not at least partially responsible for his decision.

It is important to note, however, that Frankfurt is not claiming that "the ability to do otherwise" is always irrelevant to one's moral blameworthiness. Rather, he is claiming that the ability to do otherwise is not a *necessary* feature of moral responsibility. Frankfurtian-compatibilism can account for situations where one is forced to act in a certain way due to extenuating circumstances.⁸ On the other hand, the principle of alternate possibilities provides little use in helping the court assess the responsibility of the accused. The next thought experiment is intended to continue to pump our intuitions in the direction that alternate possibilities are unnecessary to deem one responsible for their actions.

Derivative Responsibility and Future Actions

Imagine that Alex had an incredibly strong pre-disposition and first-order desire to harm opposing fans when watching sports. Alex knows that if she went to a sports bar, it would be very likely that she would run into a fan of the opposing team. If such events took place, Alex would not be able to control herself from throwing a punch or two and defending her team's reputation. In other words, Alex could not have done otherwise than how she acted. She had no alternative than to punch the opposing team's fan. If PAP is true, then it appears she is not morally responsible for her actions, as she could not have done otherwise. However, if one applies Frankfurtian-compatibilism, Alex would be held morally responsible because a first-order desire to punch an opposing fan is contrary to moral standards. In addition, it is also true that Alex could've prevented the situation by avoiding situations where she would harm others.

8. One can imagine a multitude of examples where persons are being coerced to act against their will. In these circumstances, by applying the L-Reply, one can look to a person's first and second order desires to see if such desires are truly deserving of praise and blame. If a person has a desire to not harm someone, but is being coerced into harming someone, then their act of harming someone can be understood as a higher order desire to avoid harm themselves. In this sense, harming another by possessing a desire to avoid harm yourself is likely not an immoral desire to possess. In this sense, one is looking at the reasons and basis for why one is acting, not simply whether they could have done otherwise.

Now imagine that Alex acknowledges her inclinations and first-order desires and therefore, decides to stay away from sporting events where it is likely she will encounter opposing fans. Instead of attending sports events, she stays at home. In addition, Alex decides to go see a counsellor to eliminate her first-order desire of harming opposing fans. After months of counselling, she comes to the realization that other fans are people just like her and shouldn't be harmed. Alex now has a strong disposition and first-order desire to not harm others. So, she goes to the sports bar and enjoys a game and even a bit of banter with opposing fans. At one point, she felt her past inclination to harm opposing fans, but was reminded by her counselling sessions, and couldn't harm them. In this sense, she *can't* harm them, due to her new insight and respect for others. In these circumstances, PAP would arrive at the result that Alex is not responsible for not harming the other fans, as she couldn't have done otherwise. However, the L-reply coupled with Frankfurtian-compatibilism, can praise Alex, as it acknowledges her reasons for not harming the fans and taking the appropriate steps to mitigate her threat.⁹ This thought experiment is intended to target our intuitions and ideas about what it means to be responsible from the point of view of derivative responsibility. Derivative responsibility refers to "cases where the agent is said to be blameworthy for what she did (or failed to do) by virtue of being blameworthy for the causal conditions that led to it" (Mcenna and Widerker, 2018, 14). As we saw in both scenarios featuring Alex, the conditions that led to her harming others and the conditions that led to her attending counselling and not harming others were ostensibly within her control. Therefore, appealing to our intuitions as well as the principles at play in the criminal law, Alex appears to be (derivatively) morally responsible, a feature of responsibility that seems to be worth preserving.

The determinist objection that voluntary actions require the ability to do otherwise remains unconvincing. The morally salient features of decision making can be present even if one cannot act in a way other than how one did. Frankfurtian-compatibilism (as well as the L-reply) looks to the agent's basis and reasons for acting in a certain way by examining one's hierarchy of desires. The next objection to Frankfurtian-compatibilism will challenge whether persons can be the appropriate source of their actions when the assumption of determinism is taken into account.

9. This example is based on the examples presented in Dennett, D. C. (2015). *Elbow Room: The Varieties of Free Will Worth Wanting* (New ed.). MIT Press at page 147.

Manipulation Arguments

Manipulation arguments intend to demonstrate that persons do not possess the appropriate type of control over their actions to be held morally responsible. Put differently, persons cannot be considered the proper *source* of their actions. As Caruso writes, “The basic idea behind this argument is that if an agent...is causally determined to act in a particular way...then the agent is *intuitively* not morally responsible for that action, even if they satisfy all the prominent compatibilist conditions¹⁰ on moral responsibility” (Caruso and Dennett, 2021, 53). For those persuaded by manipulation arguments, the conditions listed above in the criminal justice system (i.e., voluntariness, rationality, autonomy, lack of external forces etc.) are insufficient to demonstrate that any person could properly be considered a candidate deserving of praise or blame.

Manipulation arguments commonly invoke imaginary malevolent agents that attempt to control a subject. After such a scenario is provided, an analogy is made between the malevolent agent and the natural causal processes that are implied by determinism. Manipulation arguments are similar to the thought experiments utilized in the previous section, as they are intended to pump and move our intuition to conclude that persons should not be held morally responsible for their actions. Pereboom¹¹ offers the following manipulation argument:

Scenario 1: A team of neuroscientists has the ability to manipulate Plum’s neural states at any time by radio-like technology. In this particular case, they do so by pressing a button just before he begins to reason about his situation, which they know will produce in him a neural state that realizes a strongly egoistic reasoning process, which the neuroscientists know will deterministically result in his decision to kill White. Plum would not have killed White had the neuroscientists not intervened, since his reasoning would then not have been sufficiently egoistic to produce this decision. But at the same time, Plum’s effective first-order desire to kill White conforms to his second-order desires. In addition, his

10. These conditions would include that her mental states are causally efficacious, that she approves of her own behaviour, that she is responsible to reasons, and that she satisfies all relevant senses of control over her actions.

11. The original manipulation argument present’s four cases. However, for our purposes here, my responses apply equally as well to the two-scenario argument, as to the four-scenario argument.

process of deliberation from which the decision results is reasons-responsive; in particular, this type of process would have resulted in Plum's refraining from deciding to kill White in certain situations in which his reasons were different. His reasoning is consistent with his character because it is frequently egoistic and sometimes strongly so. Still, it is not in general exclusively egoistic, because he sometimes successfully regulates his behavior by moral reasons, especially when the egoistic reasons are relatively weak. Plum is also not constrained to act as he does, for he does not act because of an irresistible desire – the neuroscientists do not induce a desire of this sort (2014, 79).

As one can see in the first scenario, Plum is manipulated by the neuroscientists, who produced in him mental states that lead to his decision to kill White. In scenario one, the reader is supposed to conclude that Plum is not responsible for his actions, despite Plum's apparent satisfaction of acting on first and second-order desires. In addition, Plum is also reasons-responsive, acting within his own character, under no direct external constraint, and he could've done otherwise (as the desire was not irresistible). Despite these considerations, however, it still seems that Plum is not responsible. If one considers Plum to not be morally responsible in scenario one, then according to Pereboom, Caruso and others, one should also deem Plum to not be morally responsible in scenario two:

Scenario 2: Everything that happens in our universe is causally determined by virtue of its past states together with the laws of nature. Plum is an ordinary human being, raised in normal circumstances, and again his reasoning processes are frequently but not exclusively egoistic, and sometimes strongly so (as in Cases 1). His decision to kill White issues from his strongly egoistic but reasons-responsive process of deliberation, and he has the specified first and second-order desires. The neural realization of Plum's reasoning process and decision is exactly as it is in Cases 1; he has the general ability to grasp, apply, and regulate his actions by moral reasons, and it is not because of an irresistible desire that he decides to kill (Pereboom, 2014, 78).

compos mentis

The second scenario is almost exactly like scenario one, however, the egotistic desires that Plum possesses are “natural” compared to scenario one, where they were implanted by the neuroscientists. If one applies these two scenarios to Frankfurt’s conception of free will, it seems difficult to understand how Plum can be responsible for acting in line with his desires if such desires were not a result of him choosing those desires. According to Frankfurt, freedom consists in being “free to will what [one] wants to will” (1971, 16). However, in scenario one and two, Plum’s will had either been manufactured within him or he was born with it. In both cases, it seems counterintuitive to claim that freedom consists of acting in accordance with one’s will when we are simply given a will, without much say in the matter. Therefore, if Plum is not responsible in scenario one, Plum also appears to not be responsible in scenario two. By concluding that Plum is not responsible in scenario two, one is supposed to also conclude that persons are never morally responsible, as the conditions in scenario two are the conditions of the actual world we inhabit.

In response to manipulation arguments, there are two general approaches. The first set of approaches are known as hard-line replies. Hard-line replies focus on the first scenario and argue that despite the direct manipulation, Plum would still be morally responsible for his actions. The second set of approaches are known as soft-line replies. Soft-line replies attempt to articulate a principled distinction between the first and second scenario. If a principled distinction can be demonstrated, the move from scenario one to scenario two becomes suspect, thereby, allowing one to retain Plum’s responsibility in scenario two. In what follows, I will present two responses, a hard-line reply as well as a soft-line reply.

A Hard-Line Reply

Regarding hard-line responses, Caruso writes, “The first problem with the hard-line approach is that it conflicts too deeply with our intuitions about source hood” (2021, 78). It is hard to dispute Caruso’s claim that attributing responsibility to Plum in scenario one does not conflict with our intuitions. Plum is clearly manipulated and is unaware of such manipulation, which appears to negate his responsibility either partially or fully. If such a case were to be presented before a court, Plum would likely be acquitted of the murder. However, it is interesting to consider whether or not Plum would be morally responsible if he *knew* that he had been previously manipulated. Does *knowledge* of the manipulation change Plum’s

responsibility? If one knows that they possess a first and second-order desire that may lead to murder, are they not able to have another second-order desire (or even a third-order desire) that would counter-act the harmful desires? It seems reasonable, or at least possible, that in learning that someone has manipulated you to be more egotistical, one can take appropriate steps to either eliminate that desire by acquiring a higher desire or by mitigating the problematic desire.

In a similar vein to the thought experiment where Alex possesses an inclination to hurt opposing fans, when one is aware of a potential loss of control, there is a responsibility to reasonably prevent such a circumstance from occurring. In Frankfurtian terms, if one is aware of a problematic first-order desire, one should take reasonable steps to either eliminate that desire or replace it with a good desire. This sort of responsibility (i.e. taking responsibility to prevent future actions), as mentioned above is called *derivative responsibility*. There are differences between Plum's and Alex's scenarios, however, the general principle holds. That principle being, that persons are responsible for taking reasonable precautions and steps to avoid potential harm to others. Dennett calls this sort of planning, "meta level control thinking" (2015, 73) and Fischer calls it "guidance control" (2012, 178-205). So, if Plum was not directly responsible for killing White because he was manipulated, he may still be derivatively responsible for killing White, as he failed to take the appropriate steps to mitigate the manipulation.

The court recognizes this sort of control and responsibility over our future actions. For example, in contrast to the subjective mens rea (ex. intention), there is also an objective standard of mens rea. When the court applies an objective standard of mens rea, it assesses whether a reasonable person in the scenario that the accused found themselves in, would have acted different? For example, in cases where one causes the death of another while driving, the standard applied is that of a *marked departure*. The court will take into consideration all the surrounding circumstances to determine if the accused acted in a way that constitutes a "marked departure" from what a reasonable person in the same circumstances would have done.¹² Momentary lapses of judgment or accidents do not constitute a marked departure of the reasonable person.¹³ For example, an unexplained swerve into traffic or the inability to control a car in snow both constitute momentary lapses of judgment that do not incur criminal responsibility.

12. *R v Hundal*, [1993] 1 SCR 867.

13. *R v Roy*, 2012 SCC 26 and *R v Beatty*, 2008 SCC 5.

compos mentis

In this regard, there are existing legal tools and concepts that can be applied to cases of manipulation which would help the courts determine if a *reasonable* person in Plum's situation would have taken steps to prevent the murder of White. If Plum knew that he was manipulated and had the time to take steps to prevent the murder but didn't, such factors lean towards Plum being held partially responsible. However, if Plum had no idea that such manipulation took place, it seems that he would not be held morally responsible.

In response to Caruso's claim that manipulation arguments demonstrate that the subject is intuitively not responsible, Dennett writes, "I *am* responsible for my *abstention* whether or not the manipulation I have ordered is required. (Maybe you disagree, but I think you have to admit that it is far from obvious in my example that I am *not* responsible because I have been manipulated, which is the "obvious" intuition the other examples are supposed to pump)" (Caruso and Dennett, 2021, 59). In Dennett's mind, once Plum becomes aware of the manipulation, such manipulation ceases to be able to control Plum's ability to resist the egotistic desires, as Plum now can take steps to prevent his killing of White. Therefore, it seems that hard-line replies seriously question the moral intuition that is supposed to arise in us, once we consider the subject possessing knowledge about the manipulation of their own mental states and desires.

A Soft-Line Reply

The soft-line response that will be presented will further dissect the manipulation argument's presuppositions regarding the level of control needed by an agent in order to be held morally responsible. If we grant that Plum is not morally responsible for White's death in scenario one, does that require us to also conclude that Plum is not responsible in scenario two? There seems to be a few reasons to think otherwise.

The principled difference between the two scenarios is that the manipulation in scenario one is performed by conscious agents, compared to agentless physical causes in scenario two. It seems to be far more intuitive to conclude that there is a difference between agent manipulation and agentless manipulation, then to conclude that these two sorts of manipulation are equivalent. One should even consider whether manipulation is the correct term to refer to the natural causal influences that people experience. Caruso writes, "Softline replies are therefore unconvincing because, at best, they can only show that a *particular manipulation*

example has failed to capture all the relevant compatibilist conditions for moral responsibility, not that manipulation arguments fail “tout court” (2021, 84). Stated differently, Caruso appears to claim that one can simply manufacture a new manipulation argument that would demonstrate how one lacks moral responsibility. This might be true, but the burden of proof remains on the one convinced of the manipulation argument to think of such an example. It seems incredibly difficult to bridge the divide between agents being able to control agents, and the physical universe being able to control agents.

For example, as mentioned in the previous section, the court will not hold someone criminally responsible for a “momentary lapse of judgment” when driving their car. Persons can control their car, by speeding up and slowing down as well as a variety of other controls. However, when it is snowy, as was the case in *R v Roy*,¹⁴ our control of our vehicles is drastically reduced because of other causal factors. Other times, people have more control over their vehicles. For example, think of a Nascar driver who can control their vehicle far better than the average person. It seems intuitive that within driving a car, there are degrees of respective freedom that can and often are influenced by causal influences.

In the same way, we as persons are influenced by a variety of causal factors including, environment, parents, genes, predispositions etc. People also face different restrictions on their physical and mental capacities. However, despite all these limitations on our control of ourselves, it does not seem reasonable to conclude that we are not morally responsible for our actions.¹⁵ Given our understanding about control and the variations of control, we can clearly see the difference between an agent manipulating our car (via a remote control) thereby, negating our control, compared to physical processes that simply *limit* our control. Therefore, the analogy between scenario one and scenario two seems to fail due to their being a principled difference between the two scenarios. Namely, the world and the physical causes within it are not an agent which can exert the same sort of conscious and intentional manipulation on persons as a neurosurgeon can. In this sense, persons seem to possess the appropriate amount of control as well as remain an appropriate source, to be considered morally responsible for their actions. The analogy that is supposed to be drawn between scenario one and

14. *R v Roy*, 2012 SCC 26.

15. This portion relies heavily on Dennett’s account of control in his *Elbow Room*, chapter 3: Control and Self-Control.

compos mentis

scenario two is weak and thereby, should not lead one to think that persons are incapable of being morally responsible.

CONCLUSION

As mentioned in section one, criminal law presupposes legal personhood, autonomy and rationality. These three elements ground the criminal law's practice of apportioning responsibility to person's accused of committing crimes. However, according to free will skeptics, there are good reasons to deny that persons should be held morally responsible for their actions. In response to such a radical claim, Frankfurtian-compatibilism was presented as an alternative approach intended to demonstrate that the practice of apportioning responsibility is philosophically defensible. Two primary objections (i.e. the principle of alternate possibilities and manipulation arguments) that are commonly raised against compatibilism were shown to be unsuccessful or at the very least, inconclusive. As Morse writes, "Compatibilism is the only metaphysical position that is consistent with both the criminal law's robust conception of responsibility and with the contemporary scientific worldview" (2013, 123). Therefore, it seems to me that the criminal justice system's practice of holding persons responsible for their crimes can be justified on a Frankfurtian-compatibilist basis and is not be threatened by the arguments put forth by free will skeptics.

REFERENCES

- Caruso, Gregg, and Owen Flanagan. 2018. *Neuroexistentialism: Meaning, Morals, and Purpose in the Age of Neuroscience*. New York: Oxford University Press.
- Caruso, Gregg. 2021. *Rejecting Retributivism: Free Will, Punishment, and Criminal Justice*. New York: Cambridge University Press.
- Caruso, Gregg, and Daniel Dennett. 2021. *Just Deserts: Debating Free Will*. New York: Polity.
- Dennett, Daniel. 2015. *Elbow Room: The Varieties of Free Will Worth Wanting* (New ed.). Cambridge: MIT Press.

- Ferguson, Gerry. 1989. "Critique of Proposals to Reform the Insanity Defence". *Queen's Law Journal* 14 (1): 135-152.
- Fischer, John Martin. 2012. *Deep Control: Essays on Free Will and Value*. New York: Oxford University Press.
- Frankfurt, Harry. 1969. "Alternate Possibilities and Moral Responsibility." *The Journal of Philosophy* 66 (23).
- Frankfurt, Harry. 1971. "Freedom of the Will and the Concept of a Person." *The Journal of Philosophy* 68 (1), 5-20.
- Haji, I. 2012. "Reason, Responsibility, and Free Will: Reply to My Critics." *The Journal of Ethics* 16 (2): 175-209.
- Kane, R. 2005. *The Oxford Handbook of Free Will* (1 ed.). New York: Oxford University Press.
- McKenna, Michael and Justin Coates. 2021. "Compatibilism." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta.
- McKenna, Michael, and Derk Pereboom. 2014. *Free Will: A Contemporary Introduction*. New York: Routledge.
- McKenna, Michael, and David Widerker. 2018. *Moral responsibility and Alternative Possibilities: Essays on the Importance of Alternative Possibilities* (First ed.). New York: Taylor and Francis.
- Morse, Stephen and Thomas Nadelhoffer. 2013. *The Future of Punishment*. New York: Oxford University Press.
- Naffine, Ngaire. 2009. *Law's Meaning of Life: Philosophy, Religion, Darwin, and the Legal Person*. 1st ed. Portland, Or; Oxford: Hart Pub.
- Nedelsky, Jennifer. 2011;2012;. *Law's Relations: A Relational Theory of Self, Autonomy, and Law*. Oxford; New York: Oxford University Press.
- Pereboom, Derk. 2014. *Free Will, Agency, and Meaning in Life*. New York: Oxford University Press.