

July 2017

Free Will, Determinism, and Moral Responsibility: An Analysis of Event-Causal Incompatibilism

Gunnar Footh

Macalester College, gfooth@macalester.edu

Follow this and additional works at: http://digitalcommons.macalester.edu/phil_honors



Part of the [Philosophy Commons](#)

Recommended Citation

Footh, Gunnar, "Free Will, Determinism, and Moral Responsibility: An Analysis of Event-Causal Incompatibilism" (2017). *Philosophy Honors Projects*. 10.

http://digitalcommons.macalester.edu/phil_honors/10

This Honors Project is brought to you for free and open access by the Philosophy Department at DigitalCommons@Macalester College. It has been accepted for inclusion in Philosophy Honors Projects by an authorized administrator of DigitalCommons@Macalester College. For more information, please contact scholarpub@macalester.edu.

Free Will, Determinism, and Moral Responsibility:
An Analysis of Event-Causal Incompatibilism
By Gunnar John Footh

Professor Geoffrey Gorham
Department of Philosophy
4/26/17

Chapter 1: Introduction

The question of moral responsibility has been around for millenia. What is moral responsibility? How do we define it? What does it mean to be a moral person? This philosophical topic alone has been analyzed and debated among philosophers for centuries. The very existence of the debates over moral responsibility and value theory as a whole evidences the importance humanity puts on answering these moral questions. It is no surprise that these topics are still being discussed and debated today.

The moral responsibility-determinism debate is ongoing in contemporary philosophy, and it asks the following question: is moral responsibility reconcilable or compatible with a deterministic universe? Here I will define determinism as *causal determinism*, the idea that every event is necessitated by antecedent events in conjunction with the laws of nature. By the *laws of nature*, I mean those laws that govern, inform, and infringe upon our universe, the laws of physics and mathematics.

Historically, the moral responsibility-determinism debate has been focused on the following question: are alternative possibilities required for moral responsibility? Many philosophers thought so, since they connected alternative possibilities necessarily with the existence of free will. Other philosophers rejected the need for alternative possibilities but still tried to save free will from the imposing laws of nature by constructing compatibilist accounts of free will and determinism. Still others rejected moral responsibility and free will altogether. For a time it seemed that the debate between these two camps was at a standstill.

Some of these debates are still ongoing, but a shift has occurred in the moral

responsibility-determinism debate. Philosopher Harry Frankfurt in 1969 published an influential article which sought to debunk the notion that alternative possibilities were required for moral responsibility. These counterexamples, coined “Frankfurt-style” cases, attempted to show that agents could be morally responsible for their actions in a wide range of cases in which they had no alternative possibilities. These cases alone ushered in a whole new area of philosophy in the debate to challenge the rhetoric of the opposing camp, the compatibilists and incompatibilists firing counterexample upon counterexample at each other for half a century.

More recently a new philosophical theory has taken hold: Van Inwagen’s Direct Argument. Contemporary philosopher Peter Van Inwagen sought to show, using modal logic, that, since no moral agent could be responsible for past events by antecedent events and the laws of nature, via determinism no moral agent could be responsible for states of affairs that presently obtain. This theory seemed to sidestep the debate over alternative possibilities altogether, but many other contemporary theorists have had their doubts.

To analyze all of these competing theories would require a work much longer and more detailed than this. In this project, I will analyze, summarize, and critique the incompatibilist theory known as source incompatibilism, which argues that a moral agent is morally responsible for an action only if they are the proper source of that action. More specifically, I will analyze the source incompatibilist views of event-causal incompatibilism, which argues that an agent has free will only if there exists indeterminacy in her decision-making process, either before the formation of a decision itself or during the formation of a decision. I will argue that event-causal incompatibilist

views suffer from problems of control and moral chanciness. Thus I will argue that event-causal incompatibilism is no more philosophically tenable than its compatibilist counterparts. If this is true, the event-causal incompatibilist ought to abandon it due to considerations of parsimony.

After I have successfully refuted event-causal incompatibilism, I will introduce a novel theory of moral responsibility compatibilism of my own, which I will argue is the only tenable philosophical theory left for the proponent of event-causal incompatibilism. I will attempt to reconcile moral responsibility with causal determinism, utilizing an argument from the philosophy of David Enoch in his book *Taking Morality Seriously*. When this is complete, I will defend my compatibilist theory from various objections by philosophers Saul Smilansky and Ishtiyaque Haji.

I will end the discussion with a brief introduction to other non-libertarian views of moral responsibility and determinism, which do not require libertarian notions of free will and thus do not require indeterminacy for freedom. These include Saul Smilansky's illusionism and Derk Pereboom's hard incompatibilism. I will analyze these views, but ultimately I will critique them. I will argue that these theories also are lacking, and so they are not viable alternatives to the proponent of moral responsibility.

Chapter 2: Responsibility Incompatibilism Analysis

2.1 Introduction

The purpose of this chapter is to elucidate different views of moral responsibility, namely nonaction-centered and action-centered event-causal incompatibilist views, and analyze

them. In general, incompatibilists argue that incompatibilism is the best theory we have because it does not suffer from problems of moral chanciness, and it offers agents more control over their actions than compatibilist theories.¹ However, I will argue that event-causal incompatibilist views also fall victim to problems of moral chanciness and control. Thus I will argue that event-causal incompatibilism is no more philosophically tenable than compatibilism. If this is true, I will argue that event-causal incompatibilists should reject incompatibilism and adopt compatibilism for the sake of parsimony. This conclusion will bridge the gap between this chapter and the next one, in which I will explicate a compatibilist theory of moral responsibility of my own.

Responsibility incompatibilism is the view that, for any possible world in which causal determinism is true, if an agent *A* performs an action *c*, *A* is not morally responsible for *c*; there is no possible world for which causal determinism is true and an agent *A* is morally responsible for *c*. There exist two fundamentally different types of incompatibilism: leeway incompatibilism and source incompatibilism.² Both leeway incompatibilism and source incompatibilism agree that the truth of determinism is sufficient for the nonexistence of moral responsibility. Nevertheless, these two theories differ in terms of what each theory claims is necessary for moral responsibility.

Leeway incompatibilism claims that alternative possibilities are necessary for moral responsibility. These alternative possibilities allow for free will and moral responsibility, where our free will consists having metaphysically-available alternatives

¹ By “moral chanciness” I mean (in general terms) situations in which it is a matter of luck or chance that one person performs a certain moral action.

² Timpe, “Source Incompatibilism and Its Alternatives,” 143. Kevin Timpe offers a comprehensive discussion on the basic components of these different incompatibilist views.

to action. We may call this the Principle of Alternate Possibilities:

PAP: A moral agent is responsible for a state of affairs only if that state of affairs obtains and the moral agent could have done otherwise.³

Leeway incompatibilism is attractive for two main reasons: it is simple and intuitive. For proponents of *PAP*, the freedom to choose consists in an agent having genuine alternatives. For example, if given the option between lying to a friend or telling the truth, an agent has genuine alternatives just in case she could at that time have performed any of the two (or more) choices before her.

In 1969 Harry Frankfurt published a provocative article titled “Alternative Possibilities and Moral Responsibility” that challenged *PAP*. Frankfurt argued that the problem with *PAP* is that it argues that alternative possibilities are required for moral responsibility, but there seem to exist cases which contradict this claim:

It asserts that a person bears no moral responsibility--that is, he is to be excused-for having performed an action if there were circumstances that made it impossible for him to avoid performing it. But there may be circumstances that make it impossible for a person to avoid performing some action without those circumstances in any way bringing it about that he performs that action.⁴

Thus Frankfurt argued the existence of alternative possibilities is not required for moral responsibility. To better understand Frankfurt’s argument, consider the following case taken from the same article:

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 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

³ Ibid. 143.

⁴ Frankfurt, “Alternative Possibilities and Moral Responsibility,” 837.

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' initial preferences and inclinations, then, Black will have his way.⁵

Cases like these are known as “Frankfurt-style” cases, and they show up frequently in the literature concerning *PAP*. What Frankfurt attempted to show was that we often regard agents (i.e. Jones, in this case) as morally responsible for their actions even without alternative possibilities. Whether the agent *believes* they are the proper source of their actions is moot. *PAP* is false.

I believe Frankfurt is successful in debunking *PAP*. While other forms of *PAP* have been developed in the literature (like Van Inwagen's Principle of Possible Prevention (*PPP*)⁶), it is dubious what advantages these have over their *PAP* counterparts. The burden on the leeway incompatibilist is to explain why alternative possibilities are required for moral responsibility at all. One attempt to fix leeway incompatibilism is to assert that there exists a more basic, fundamental requirement for moral responsibility. This fundamental requirement is that moral agents are the whole sources of their moral actions, which is the view of source incompatibilism.

Source incompatibilism claims that an agent being the original source of her actions is a requirement for moral responsibility.⁷ An agent being the proper source of her

⁵ Ibid. 835.

⁶ Van Inwagen, “Moral Responsibility, Determinism, and the Ability to do Otherwise,” 345. *PPP* fails because there seem to be cases in which we hold agents morally responsible for states of affairs that could not have been prevented. In the case of suicide, one might hold me morally responsible for an agent's suicide, even if it was determined to happen, if I had not done ‘all I could’ to prevent it, it being my duty of course to prevent such things in the first place.

⁷ Timpe, “Source Incompatibilism and Its Alternatives,” 143.

actions substantiates free will and allows for moral responsibility. We may call this the Principle of Agent Ultimacy:

PAU: A moral agent is responsible for a state of affairs only if that state of affairs obtains and the state of affairs can be traced back properly to the moral agent.

The point here is that the moral agent themselves generate the action in question. The difference between source incompatibilists and leeway incompatibilists is that source incompatibilists are not required to embrace *PAP*. For source incompatibilists, moral responsibility does not require merely alternative possibilities. Whether or not the agent is the source of her moral actions determines moral responsibility, where the source herself is not determined by external factors. What it means for an agent to be the proper source of her actions varies among source incompatibilist theories.

The two main theories of source incompatibilism are agent-causal and event-causal views.⁸ *Agent-causal* views argue that free actions must be caused by an agent, and neither what the agent causes to happen nor the agent's causing something to happen is determined by prior events.⁹ In agent-causal views, the agent is a persisting substance, which itself cannot be an effect, and free action is generated by this substance. Thus agent-causal views then emphasize the existence of an agent-substance that persists and acts, and they require the falsity of causal determinism.

Event-causal views argue that free actions must be caused by an agent, and this causation consists in indeterministically-caused agent-involving events.¹⁰ In event-causal

⁸ Clarke, Randolph and Capes, Justin, *Incompatibilist (Nondeterministic) Theories of Free Will*, Stanford Encyclopedia of Philosophy. This source was invaluable for my research and the writing of this project.

⁹ Ibid.

¹⁰ Ibid.

views, free action is generated by an agent exercising some sort of causal control that is consistent with determinism. However, event-causal views emphasize that actions are free just in case that there is indeterminism somewhere in the causal chain of the production and employment of an agent performing an action.

The account I wish to target in this essay is event-causal incompatibilism. Traditionally event-causal incompatibilist accounts of moral responsibility are built off of compatibilist accounts of free will and determinism. For the purpose of this essay however, the fact that these accounts feature indeterminism marks them for critique. *Event-causal incompatibilism* is the view that an agent is only morally responsible for their actions if indeterminism is present in the production of the agent-involving events.¹¹ Where this indeterminism occurs is a matter of debate among event-causal theorists.

In this essay I consider two possible event-causal views: nonaction-centered and action-centered. *Nonaction-centered* event-causal views locate the indeterminism early in the causal chain before the formation of the agent's decision. *Action-centered* event-causal views locate the indeterminism at the precise moment in the causal chain of the formation of the agent's decision. I will analyze these two views of event-causal incompatibilism and identify two fatal problems. One problem is based on moral chanciness; the other concerns lack of proximal control. I will argue that, due to these problems, event-causal incompatibilism falls prey to the same arguments against compatibilist theories of moral responsibility and free will. So event-causal incompatibilist views are no more tenable than their compatibilist counterparts.

¹¹ Ibid.

2.2 Brief Analysis of Control and Action

Philosophical theories concerning free will and moral responsibility place emphasis on describing and analyzing the kind of control necessary for agents to be responsible for moral actions. Ishtiyaque Haji provides a good summary of the kind of control associated with event-causal incompatibilist views of moral responsibility and determinism. In order to make sense of nonaction-centered and action-centered event-causal views, a summary and analysis of this type of control is required, which I will explicate here.

The control in question is proximal control. Haji notes that proximal control concerns the direct causal production of agent-involving events. These include but are not limited to: an agent's having certain values, desires, and beliefs; an agent making a certain evaluative judgment; an agent forming a certain intention or decision; an agent executing an intention; or an agent performing a nonmental action.¹² Any type of proximal control is free from influences that would completely undermine the agent's freedom, influences like compulsion, manipulation, and insanity according to philosopher Alfred Mele.¹³

By definition, all physical or mental occurrences by an individual are events, where an action is an exercise of some sort of direct (usually conscious) control by an

¹² Haji, "Alternative Possibilities, Luck, and Moral Responsibility," 257. While technical, Haji's treatment of proximal control and moral luck concerning nonaction-centered and action-centered event-causal incompatibilist views form the backbone of my objections in this honors project. This essay is a landmark of the case against event-causal views, in my humble opinion.

¹³ Mele, *Autonomous Agents: From Self-Control to Autonomy*, 222. I must note here that Mele's nonaction-centered view is not a view that I am ascribing to him. Mele is agnostic about the free will and moral responsibility debate. He offers merely a proposal for nonaction-centered event-causal incompatibilism.

agent.¹⁴ An agent is said to have direct actional control in some situation if an agent performs an action whose cause is a direct function of the agent's character and will - her desires, beliefs, intentions, etc. For example, imagine an agent, Marie, sees a wallet on the ground. In light of her values, desires, and beliefs, Marie forms the intention to give the wallet back to its owner. After all, she does not need the money, and she wants to do the moral thing. Here Marie is exercising direct actional control, because she performed an action (i.e. forming the intention to give the wallet back to its owner). More generally, Marie is exercising proximal control, because she was involved in the direct causal production of an agent-involving event (i.e. forming an intention at all).

Agents can also exercise indirect actional control. This occurs when an agent exercises control over the occurrence of an event, but this control is derived from the agent exercising direct actional control over some earlier action. For example, imagine an agent Bob who sees a child drop her ice cream cone on the ground. Bob sees the child start to cry, and this resonates with him. Bob forms the evaluative judgment that helping the child would be a good thing. Here Bob is exercising indirect actional control. He formed an evaluative judgment which is not an action by definition, but this judgment is a function of Bob's earlier intention (let's say) to be kind to others. Bob is exercising proximal control because he too was involved in the direct causal production of an agent-involving event (i.e. the formation of an evaluative judgment).

Proximal control can even take a non-actional form. This occurs when an agent exercises control over an event that is not an action. For example, imagine an agent Sally

¹⁴ Haji, "Alternative Possibilities, Luck, and Moral Responsibility," 257.

who goes to church for the first time. In church, Sally comes to value the atmosphere and effects of worship. While Sally has no direct control over this event, she still has proximal control. because she is involved in the causal production of an agent-involving event (i.e. coming to value something). It is important to note that having proximal control does not require previously having direct or indirect actional control over an agent-involving event.

In short, event-causal incompatibilism postulates that the control required for free action and moral responsibility is a kind of causal control.¹⁵ Furthermore, event-causal incompatibilism requires that choices, decisions, or intentions for which agents are morally responsible be outcomes of causal processes.¹⁶ Event-causal views argue that, in order for agents to act freely and responsibly, they must have the capacity to engage in causal control and practical reasoning to guide their behavior in light of the reasons they have for acting. This requirement is known as *reasons responsiveness*, in which agents are responsive to reasons which may or may not influence their actions, depending on the strength of the reasons in deliberation. However, in contrast, the agent's free decision is in part indeterministically caused.

The kind of control necessary for moral responsibility is not merely proximal control but what is known as "ultimate control."¹⁷ Thus event-causal incompatibilists (and libertarians in general) argue in turn that ultimate control is a requirement for moral responsibility, and that ultimate control is only possible if determinism is false. If the

¹⁵ Ibid. 255.

¹⁶ Ibid. 256.

¹⁷ Clarke, Randolph and Capes, Justin, *Incompatibilist (Nondeterministic) Theories of Free Will*, Stanford Encyclopedia of Philosophy.

incompatibilist theory is more philosophically tenable than compatibilism, it is because it offers us this ultimate control which is precluded by compatibilist theories. What exactly ultimate control amounts to depends on whether the theory itself is nonaction-centered or action-centered. I will consider these two event-causal views in the next sections.

2.3 Nonaction-centered Event-Causal Incompatibilism

Nonaction-centered event-causal incompatibilism is the view that indeterminism is located early in the causal pathway of an agent performing an action. More specifically, this indeterminacy lies in the region of the causal pathway *before* an agent makes a decision, and the indeterminacy is not caused by actions of any sort.¹⁸ However, the causal pathway from the formation of one's best judgment to performing an action is deterministic.

To help illuminate this view, consider the following example. Susan is deciding whether or not to cheat on her physics exam. When Susan engages in this moral deliberation, a number of things happen. Before she forms her best judgment about what to do, a number of considerations come to mind during her deliberations. Perhaps Susan remembers that she needs to pass this exam to pass this class. Perhaps also that Susan remembers her Christian upbringing, and she feels that cheating would be the morally wrong action to commit. Perhaps Susan thinks that no one is watching, and she can get away with cheating. All of these mental states - beliefs, values, reasons, etc. - flood into Susan's mind. She considers each mental state but ultimately decides to cheat.

¹⁸ Ibid.

The nonaction-centered view argues that Susan is free just in case the mental states that generate her decision to cheat arise indeterministically. Nothing causes the mental objects that Susan takes under consideration to enter into her brain. Susan considers these mental objects in her deliberation, and once she has completed this deliberation she forms her best judgment about what to do. This judgment is also formed indeterministically, and under the nonaction-centered view is not an action of the moral agent at all. Once Susan's best judgment about what to do is formed, this judgment then deterministically causes Susan to form the intention and make the decision to cheat.

Another way to conceive of the nonaction-centered view is by analogy. Imagine you have a mathematical function $f(x)$. By definition, a function has a unique output for each input fed into it. Imagine however that someone gives you a random value for x . You cannot be sure what it is, but you can be sure that it will generate a unique output when fed into your function. We can draw a parallel between the nonaction-centered view and the mathematical function: the value for x is analogous to the mental objects that pop into Susan's brain before she makes her best judgment. She does not know what these mental objects will be, but she can be sure that they will generate a unique action. Here the function is analogous to Susan's decision-making process. Once fed certain initial conditions, the decision-making process begins and will deterministically spit out a unique action (i.e. a determined decision).

Philosopher Alfred Mele offers an extensive nonaction-centered account of event-causal incompatibilism in his books *Autonomous Agents* and *Free Will and Luck*. Mele begins his discussion of the nonaction-centered view with a distinction between an

agent having proximal control over an action x and ultimate control. Mele argues an agent, under a nonaction-centered view, has ultimate control over x only if there exist no conditions external to the agent that are causally sufficient for the agent performing x .¹⁹ In short, ultimate control for Mele requires the absence of determinism. Proximal control however is compatible with determinism on Mele's account, because being involved in the production of an agent-involving event requires no indeterminacy at all.

Mele's nonaction-centered account revolves around the ability of agents to engage in what he calls "full-blown, deliberative, intentional action."²⁰ For Mele this type of action requires the following items: (1) a psychological basis for practical evaluative reasoning (including but not limited to an agent's values, desires, beliefs, habits, skills, and capacities); (2) an evaluative judgment being made on the basis of such reasoning which endorses a particular course of action; (3) an intention acquired or formed on the basis of this judgment; and (4) the existence of an action that executes this intention.²¹

For example, a parent deciding where to send their child to school is an example of Mele's full-blown, deliberative, intentional action. This action is full-blown, deliberative, and intentional because the parent has values, desires, beliefs, habits, skills, capacities, etc. that factor deterministically into this decision. One of these values might be a good education for their child; one of these desires might be to provide their child with this education, and so on and so forth. The parent makes an evaluative judgment based on these mental states (which arise indeterministically during deliberation), a judgment perhaps that one particular school is better-suited for their child than any other.

¹⁹ Mele, *Autonomous Agents: From Self-Control to Autonomy*, 211.

²⁰ *Ibid.* 13.

²¹ *Ibid.* 13.

The parent then forms an intention to send their child to this school based on this judgment, and the parent executes this intention by enrolling the child into this school.

Mele's nonaction-centered theory is long and complex, but its main tenets are easily understood. Agents who engage in evaluative reasoning in light of their own motivations, make judgments, form intentions, and act according to those intentions, are free as long as they have ultimate control (i.e. as long as the mental objects responsible for determining some agent-involving event are not causally determined). Mele locates the indeterminism in this process in the emergence of mental objects before consideration and the forming of an evaluative judgment. The formation of an intention however and the action itself are results of a deterministic process.

Nonaction-centered event-causal incompatibilism emphasizes the ability for agents to act according to their intentions, desires, and considerations, exerting ultimate control which can only exist in the absence of causal determinism. While the action of decision making is ultimately a determined process, the intuition is that agents still exercise direct actional control over forming intentions, making decisions, and generating actions. However, it is unclear how nonaction-centered views demonstrate increased proximal control for agents compared to compatibilist accounts of moral responsibility and determinism.

Consider again the example of Susan who is deciding whether or not to cheat on her physics exam. On the nonaction-centered view, intentions, desires, reasons, and motivations arise spontaneously and indeterministically in Susan's brain. Once these states are in place, Susan moves on to the next step in the decision-making process:

forming her best judgment about what to do. Everything after this step is deterministic. That is, Susan's best judgment, which results in her decision, which results in her action is a function of her intentions, desires, reasons, and motivations. On the compatibilist view, intentions, desires, reasons, and motivations arise deterministically, and everything beyond judgment formation in the decision-making process is deterministic.

It is hard to see how the decision-making process in the nonaction-centered view differs from its compatibilist competitors in a significant way. Susan exercises direct active control in both views. The only difference between the two is that in the nonaction-centered view the mental objects which arise in Susan's brain are not a function of a determined causal process, and genuine alternative possibilities exist due to indeterminism. However, Susan cannot choose to do something other than what her intentions, desires, reasons, and motivations cause her to do. What's more than this, is that Susan cannot control which intentions, desires, reasons, and motivations she has in the first place. It seems then that the proximal control that agents exercise in nonaction-centered views mirrors the proximal control that agents exercise in compatibilist views.

Indeterminacy fails to increase proximal control in the nonaction-centered view because ultimate control does not enhance proximal control. At best ultimate control allows that Susan can exercise proximal control in performing *whichever* alternative she performs given genuine alternatives, giving them no more power than in compatibilist theories.²² On the nonaction-centered view, Susan lacks the capacity to ultimately

²² Haji, "Alternative Possibilities, Luck, and Moral Responsibility," 260.

determine whether or not she will cheat, since he lacks control over which mental states arise in her brain. Thus nonaction-centered event-causal incompatibilism offers no more proximal control than its compatibilist competitors.

Mele agrees with this analysis, but argues that only indeterminism can allow for agent ultimacy. It is of course up to the incompatibilist to assert *per se* that the absence of determinism allows for ultimate control, agent ultimacy, and thus moral responsibility, but this assertion is no less problematic. Certainly we are endowed with just as much power or skill to determine our futures in a deterministic universe as we are in the nonaction-centered view. The example and analysis involving Susan shows this.

Mele's argument that we have ultimate control simply because of the existence of indeterminism cannot account for the reasons why we need indeterminism for free will or moral responsibility in the first place. I am sympathetic with Haji when he argues that if ultimate control is to make a difference to free action or moral responsibility ascription, it must make a difference *because* it has some bearing on proximal control.²³ But Mele's nonaction-centered account fails to offer us this difference. For this reason, it seems the requirement for indeterminism is *ad hoc*. Nonaction-centered event-causal incompatibilism accounts offer us no more control to determine what we will do than their compatibilist rivals.

2.4 Action-centered Event-Causal Incompatibilism

Action-centered event-causal incompatibilism places the indeterminacy not in the

²³ Ibid. 261.

precursors for action but in the action itself. The process of coming to a decision may be deterministic but the performing of an action (i.e. forming an intention, making a decision, or following through with a decision) itself is indeterministic.²⁴ Thus action-centered views posit that the indeterminism that allows for free action and moral responsibility lies in the agent actually making the moral decision.

To help illuminate this view, consider the following example. Charles and his wife are at a track and field event. They are hungry, so Charles decides to get some snacks for them at the concession stand. After waiting in line and paying for the food, Charles begins the journey back to his seat and forms the intention to deliver the food safely to his wife. However, Charles is also a prankster. For a moment, on his way back, he considers spilling the food and drinks all over his wife's lap. However, when Charles returns to his seat, he promptly decides that pranking his wife is not only in poor taste, but it would also make his wife very angry. He ultimately decides to refrain from pranking his wife, and they continue watching the event without a problem.

The action-centered view argues that Charles is free only if his action of being kind and respectful to his wife (i.e. refraining from pulling the prank) is indeterministically caused. The indeterministically-caused event in Charles' predicament is his mental action to decide whether or not to prank his wife. The indeterminacy does not arise in which mental objects flood Charles' brain. Rather the indeterminacy lies in Charles' action itself. He could decide to pull a prank on his wife, or he could decide to refrain from doing so. This action is not deterministically caused by the mental objects

²⁴ Clarke, Randolph and Capes, Justin, *Incompatibilist (Nondeterministic) Theories of Free Will*, Stanford Encyclopedia of Philosophy.

that flood into his brain during deliberation.

We can understand this action-centered view also by mathematical analogy. Consider the mathematical object $g[x]$. When fed a value x , $g[x]$ spits out an indeterministic value. That is, it is not a function like $f(x)$ at all. In the action-centered view, our value x is again all those mental objects that flood Charles' brain which he used in moral deliberation to inform his decision. Charles' actually deciding to refrain from pranking his wife is analogous to $g[x]$. When fed a value x , there is no deterministic, unique output we can expect from $g[x]$. In fact, if we plug the value x into $g[x]$ in two identical possible worlds, there is no guarantee that $g[x]$ will produce the same output in both worlds.

Philosopher Randolph Clarke offers a contemporary version of action-centered event-causal incompatibilism which he calls modest libertarianism.²⁵ Clarke advocates for a theory of incompatibilism that contains all of the aforementioned qualities of the compatibilist account (those concerning reasons responsiveness, acting in light of reasons and considerations, etc.) but, in contrast to Mele's account, locates the indeterminism in the causal process of an agential decision in the direct moment of the causation of the decision itself. This exemplifies the action-centered view previously discussed.

Clarke supports a different type of ultimate control that is required for moral responsibility, which he calls "action-centered ultimate control."²⁶ For Clarke, Mele's nonaction-centered view does not allow for action-centered ultimate control. This type of control requires that there be at no time any minimally causally sufficient conditions for

²⁵ Clarke, "Libertarian Views: Critical Survey of Noncausal and Event-Causal Accounts of Free Agency," 365.

²⁶ Clarke, "Modest Libertarianism," 23.

an agent making the decision over which the agent has no direct or indirect actional control. On the action-centered view, Clarke argues that the formation of an intention and the making of a decision, being the outcome of a causal process, is the exact thing over which agents should have control if they are free, and these are lacking on the nonaction-centered view. Thus he places the indeterminism in the moment of action (i.e. making a decision itself).

For action-centered event-causal incompatibilism, an agent's previous mental states alone are not causally sufficient for the agent performing some action. Action-centered incompatibilism cites the presence of indeterminism in decision making as a potential strength, where moral agents cannot exercise the necessary agential control for moral responsibility in a deterministic universe. However, I will argue that this indeterminacy is also debilitating to the action-centered view, because it opens the theory up to problems of moral chanciness.

Consider again the case of Charles and his wife. Under the action-centered view Charles' decision to refrain from pulling the prank is nondeterministically caused, and it has no necessary bearing on his previous mental states. If this is so, there was a chance that Charles' deliberative process - the same process that lead him to the refrain from pranking his wife - could result in Charles deciding to prank his wife. Everything prior to Charles' decision might have been exactly the same, and yet he could have made the alternative decision instead. It seems dubious to argue that the action the indeterminacy results in is under Charles' control, because the only way of exercising direct actional control over an event which one is morally responsible for is by determining or

preventing it in general, and here we are concerned with the former.

If the action-centered view is correct, it does not increase the amount of proximal control that Charles has over his decision. If this is true, it is unclear then why we might regard Charles as an agent worthy of praise by refraining from pranking his wife. If his decision to refrain from pulling the prank on his wife is not causally determined by his previous mental states it does not seem to be determined by anything over which Charles has control. In any normal circumstance, we would not praise someone for doing a morally good act which is not under their control.

To help elucidate this point, consider Charles*, who did decide to prank his wife, who exists in another possible world whose past and laws of nature are identical to Charles' world: everything in Charles and Charles*' worlds up until the decision to prank his wife exactly the same. In deciding to prank his wife, Charles* originally acquires the intention to deliver the food and drinks safely to his wife for their consumption. However, when the time comes to make the decision of whether or not to prank his wife, Charles* indeterministically decides to spill the food and drinks all over her.

What's curious about this example is that there seems to be no reason why Charles* pranks his wife and why Charles refrains from doing so. The making of Charles*' decision is not explained by his prior deliberations, because these prior deliberations mirror Charles' decision to refrain from pranking his wife. Mele comments that, if one agent does one thing and another agent refrains from doing that same thing, "and there is nothing about the agents' powers, capacities, states of mind, moral character, and the like that explain this difference in outcome, then the difference really

just is a matter of luck.”²⁷

This type of moral chanciness seems incompatible with free action and moral responsibility. Haji notes that this sort of moral chanciness appears largely *because of* the availability of genuine alternatives.²⁸ The objection here is two-fold: the indeterminism present in action-centered views opens it up to cases of moral chanciness and fails to increase the proximal control an agent has in a moral situation. If this objection is successful, it not only undermines moral responsibility itself, but it also shows that the ultimate control agents have in the action-centered view offers no more proximal control over their actions than compatibilist accounts of free will and moral responsibility.

These objections against the action-centered view however have been challenged. Philosopher Randolph Clarke argues that it is not clear that the indeterminacy present in action-centered views generates control-diminishing luck at all. He says,

... Suppose that you throw a ball attempting to hit a target, which you succeeded in doing. The balls' striking the target is not itself an action, and you exercise control over this event only by way of your prior action of throwing the ball. Now suppose that, due to certain properties of the ball and the wind, the process between your releasing the ball and its striking the target is indeterministic. Indeterminism located here inhibits your success at bringing about a nonactive result that you were (freely, we may suppose) trying to bring about, and for this reason it clearly does diminish your control over the result ... But the indeterminism in [the action-centered case] ... is located differently. It is located ... in the direct causation of the decision, which is itself an action ... In the ball-throwing case, the indeterminism constitutes control-diminishing luck because it inhibits the agent from bringing about a nonactive result that she is actively trying to bring about. But that explanation is not available in the second kind of case.²⁹

²⁷ Mele, “Ultimate Responsibility and Dumb Luck,” 280. I will refer to this moral “luck” as “chanciness.”

²⁸ Haji, “Alternative Possibilities, Luck, and Moral Responsibility,” 269.

²⁹ Clarke, “Libertarian Views: Critical Survey of Noncausal and Event-Causal Accounts of Free Agency,” 365.

Clarke here is arguing that the indeterminism present in cases like the ball-throwing case are located in the causal chain *after* an action. This indeterminism constitutes control-diminishing luck because it inhibits the agent from actively bringing about the object of her action. Clarke argues that indeterminism present in cases like Charles and Charles* are located in the causal chain *during* an action. He regards the conclusion that the indeterminism present in these cases are control diminishing as inconclusive, because the intended end of the agent's action is still undergoing determination. It is not a nonactive result, like a ball striking a target, but an active result since it is produced and determined *by* an action.

However, Clarke is wrong to suppose that cases like the ball-throwing case and action-centered cases like Charles and Charles* above are dissimilar. Furthermore, I am not claiming that the indeterminism present in action-centered cases diminishes proximal control necessarily. My argument is that the proximal control we derive from cases like these do not differ at all from the control present in compatibilist accounts of moral responsibility and determinism. The moral chanciness present in these cases however is still a concern.

Consider Charles and Charles* again, living in their respective worlds w_1 and w_2 to elucidate this point. From the analysis it is clear that Charles exercises proximal control - and as indeterminism would have it - intentionally decides to refrain from pranking his wife. However, it is not up to Charles* in w_2 to determine whether or not pranks his wife. The type of indeterminism present in w_2 is the same type of indeterminism that is present in w_1 . For all intents and purposes, Charles* in w_2 engages

in exactly the same sort of reasoning that Charles engages in, in w_1 up to the point of the decision of whether or not to prank his wife. The past in both of these worlds are fixed, and they are identical up to the moment of the decision.

Consequently, nothing about Charles*'s deliberations in w_2 can explain why he decided to prank his wife, because he engaged in the same deliberations as Charles in w_1 who refrained from doing so. As Haji correctly notes, the only possible explanation for the difference between Charles and Charles*'s behavior - the "differentiating factor" - must be or involve the "indeterminacy or chanciness constitutive of nondeterministic causation."³⁰ The only possible explanation for Charles*'s deviating behavior from Charles is the indeterminism present at the moment of his decision. But certainly if this indeterminacy or chanciness is not a result of deterministic causation by the Charleses, neither Charles nor Charles* has control over this factor.

Thus the ball-throwing case and the Charles/Charles* cases are more similar than Clarke would like to admit. Clarke remarks that once you throw the ball, you have no control over the ball's trajectory. And so consequently you have no proximal control over the chanciness or indeterminism that effects ball's trajectory after it has left your hands. Throwing the ball is a basic action, the result of some intention that the agent has, in which there is indeterminism between its immediate causal antecedents and its occurrence. We can think of Charles' decision also as an outcome of prior events, the prior events being Charles' intentions and reasons for refraining to prank his wife. In this sense, there is also indeterminism between the outcome (i.e. Charles' decision to refrain

³⁰ Haji, "Alternative Possibilities, Luck, and Moral Responsibility," 268.

from pranking his wife) which is an action like throwing a ball, and its immediate causal antecedents (i.e. Charles' forming the judgment).

Haji concludes that if indeterminism in the ball-throwing case inhibits your success at bringing about the result that you were intending to bring about, which diminishes your control over the result, we should also conclude that the indeterminism in Charles*' case also diminishes the control that Charles* has in deciding as he does.³¹ For we have seen that Charles*' mental states preceding the decision included his intention to deliver the food safely to his wife for eating, but the indeterminism present in Charles*' making the decision inhibited the success of this desired result. Thus Haji argues the proximal control rebuttal is defeated.

In any case, the moral chanciness objection still has considerable power. For Clarke, chanciness detrimentally affects moral responsibility and free action then only if chanciness detrimentally affects proximal control, but I disagree. Let's suppose that any action of Charles or Charles* is an exercise of a measure of direct actional control by Charles and Charles* respectively. Haji argues that this kind of control is not sufficient for the ascription of moral responsibility. Moral responsibility ascription requires the nonexistence of what Haji calls "responsibility-undermining factors."³² These might include brainwashing, psychological manipulation, or any other phenomena that might precede or influence drastically the psychology of the moral agent. The indeterminacy in Charles*' case seems to be this sort of responsibility-undermining factor.

As Haji notes, "appraisals of responsibility are first and foremost appraisals of the

³¹ Ibid. 268.

³² Ibid. 269.

agent; they disclose the moral worth of an agent with respect to some episode in their life.”³³ When a moral agent performs a moral action, their reputation in the moral community is affected. This effect relies on the nonexistence of these responsibility-undermining factors. But clearly Charles*’ decision to prank his wife cannot reflect poorly on his moral standing if a factor of indeterminacy causally influences this decision. It does not seem to matter whether this indeterminacy is located internally or externally to the agent either.

Consider for example the case of Charles+. The case of Charles+ is exactly like Charles and Charles* up until the point at which Charles+ gets back to his seat. Upon arrival, Charles+ forms the decision to refrain from pranking his wife. However, a large gust of wind promptly blows by Charles+ and knocks the food and drinks into her lap. In the case of Charles+, our intuition is to let him off the hook. After all, the food and drinks spilling into his wife’s lap was accidental. Surely Charles+ did not intend for that to happen. The wind is an example of a responsibility-undermining factor.

Or consider for example the case of Charles=. The case of Charles= is exactly like Charles, Charles*, and Charles+ up until the point at which Charles= gets back to his seat. Charles= also forms the decision to refrain from pranking his wife. However, upon arriving at his seat, Charles= has a minor seizure, spilling the food and drinks all over his wife. In the case of Charles=, our intuition is also to let him off the hook. This internal biological phenomenon, his seizure, was also accidental. Charles= did not intend to spill the food and drinks on his wife. Charles=’ seizure here is an example of a

³³ Haji, “Alternative Possibilities, Luck, and Moral Responsibility,” 269.

responsibility-undermining factor.

If our intuitions are to let Charles+ and Charles= off the hook, they should also allow Charles* off the hook. After all, Charles* also forms the intention to refrain from pranking his wife. However, something beyond Charles*' control causes him to spill the food and drinks on his wife. Charles* did not intend for this to happen. It is just moral chanciness that he performs the moral action that he does. The burden of the incompatibilist is to explain how this indeterministic stage can be a function of any of the Charleses control, but insofar as none of their actions can be determined by their preceding mental states, it seems that none of the Charleses can determine which action the indeterminism will result in.

There seems to be no appreciable difference between the cases of Charles+, Charles=, and Charles*. If there is one, it lies in the fact that forces external to Charles+ cause his behavior, while forces internal to Charles= and Charles* cause their behavior. This however does not constitute a significant difference in our intuitions about whether or not the Charleses are responsible. Regardless of whether or not the force is external or internal to Charles, it is clear that he has no control over it, and it is this caveat that lets him off the hook. There is no good reason to let Charles= off the hook and hold any of the aforementioned Charleses as morally responsible.

One response to the objection I am making here is that there *is* an appreciable difference between the cases of Charles= and Charles*. The proponent of the action-centered case might argue that, in the case of Charles*, he himself as an agent plays a causal role in the production of the action, whereas in the case of Charles=,

something separate from him causes him to spill the food and drinks on his wife. However, this response misses the mark. It does not matter whether or not the force is external or internal to Charles= or Charles*, integrated in his agency or apart from his agency. The mistake the action-centered view makes is assuming that the agent has control *over* this indeterminism. It seems altogether impossible for an agent to have ultimate control over what happens as a result of indeterminism. The action-centered view gives us no reason to believe that control over this indeterminacy is possible or actual.

I have argued that nonaction-centered views and action-centered views suffer from two fatal objections: proximal control, and moral chanciness. Nonaction-centered views fail to offer a sense of control that is stronger than its compatibilist counterparts. Similarly, action-centered views do not increase the proximal control that an agent has over her situation. At best, they allow for the same level of proximal control that compatibilist views offer. I have also shown that the problem of moral chanciness also plagues the action-centered view. I have shown that indeterminacy is a responsibility-undermining factor, and event-causal incompatibilism can give no convincing reason why alternative possibilities *per se* is a requirement for moral responsibility, as it has no bearing on proximal control.

If my analysis is correct, there is no reason to suppose that nonaction-centered or action-centered event-causal views are more philosophically tenable than their compatibilist counterparts. Neither theory can account for “true” moral responsibility, in which compatibilist theories also (supposedly) fail. Similarly, neither theory can offer an

agent more proximal control than compatibilist theories. For these reasons, event-causal incompatibilism fails.

What is required, then, is a more robust compatibilist theory that is intuitive and answers common compatibilist objections. I offer just this theory in the next chapter. First, however, I lay out David Enoch's argument for indispensability, which grounds my compatibilist theory. In the next section, I will evince this indispensability argument, and I will offer a new compatibilist theory of moral responsibility and determinism that relies on the existence of moral facts for moral responsibility ascription. I will show that this theory stands up to various incompatibilist objections by philosophers Saul Smilansky and Ishtiyaque Haji.

Chapter 3: Moral Realism, Indispensability, and Compatibilism

3.1 Introduction

The purpose of this chapter is to elucidate and defend a theory of moral responsibility that is compatible with determinism. In the first section of this paper, I will offer a positive argument for our ontological commitments to normative facts inspired by philosopher David Enoch. I will argue that moral deliberation is an intrinsically indispensable project in which all moral agents engage. Then I will argue that this deliberation is non-optional and depends on the assumption of normative facts. Finally, I will argue that, because we cannot engage in moral deliberation without presupposing the existence of normative facts, we are justified in believing in them.

In the second and third sections, I will present a psychological analysis of moral

deliberation and reactive attitudes. This analysis will examine the different attitudes and goals that we have as moral agents when we engage in moral deliberation. My psychological analysis offers positive evidence that moral agents presume the existence of moral facts when they engage in moral deliberation. In addition I will offer an account for the moral reactive attitudes of praise and blame. Finally, I will compare my findings with the work of philosophers P. F. Strawson, David Enoch, and J. J. C. Smart in an attempt to bridge the connection between my moral-psychological analysis and metaethical argument from indispensability and moral responsibility ascription.

In the fourth section, I will bring all of these ideas together in order to form an intuitive and plausible compatibilist theory of moral responsibility. I will argue that moral responsibility follows necessarily from the existence of normative facts, and our ontological commitment to these normative facts should compel us to adopt a theory of moral responsibility that depends not on free will but adherence to these normative facts. I will defend my compatibilist theory against objections from philosophers Saul Smilansky and Ishtiyaque Haji: the argument from shallowness, the argument from complacent compliance, and the argument of moral chanciness based on the existence of possible worlds.

3.2 Enoch's Argument From Indispensability

Often philosophers of mathematics will utilize the argument from indispensability to argue for the existence of numbers. The idea is that numbers are necessary for engaging in mathematical discourse. Since we cannot refrain from using numbers in mathematical

discourse, we are justified in believing in their existence. Here, following in the footsteps of philosopher David Enoch, I will utilize the argument from indispensability to establish and justify our ontological commitment to normative facts, facts about what we *should* do. I will focus heavily on the indispensability of moral deliberation and normative facts.

In general, objects are always indispensable *to* or *for* a purpose or project. For example, mathematics is indispensable for computation, and Hawaiian shirts are indispensable for participating in Hawaiian shirt Fridays in the Macalester College physics department. Furthermore, two types of indispensability exist: instrumental indispensability and intrinsic indispensability.³⁴ I will begin the indispensability argument by outlining these two types of indispensability. Then I will argue the following: (1) normative facts are instrumentally indispensable for moral deliberation; and (2) moral deliberation is an intrinsically indispensable project.

Instrumentally indispensable objects bear a certain relationship to the project they are instrumentally indispensable *for*. Instrumentally indispensable objects are not merely useful for the project in question. For example, let's suppose money is useful for the project of purchasing a vehicle. If we can purchase a vehicle by using some other method - such as trading our old vehicle in for credit - clearly money is not instrumentally indispensable for purchasing a vehicle. This is because we can eliminate money from the project of purchasing a vehicle and still get the same results.

Similarly, merely enabling the actualization of a project does not make an object instrumentally indispensable. For example, let's suppose that a healthy lifestyle

³⁴ Enoch, *Taking Morality Seriously*, 51.

consisting of a nutritious diet and regular exercise regimen is sufficient for successfully engaging in philosophical discourse. If someone successfully engages with philosophy, we might say this is evidence of the agent living a healthy lifestyle. However, healthy lifestyles are not necessary for engaging in philosophy. Instrumentally indispensable objects bear this special relationship to the project they are instrumental for: x is instrumentally indispensable for a project or purpose y in which case x is necessary for y and cannot be eliminated without undermining our reason(s) for engaging in y in the first place.³⁵

However, instrumental indispensability alone cannot tell us which projects are intrinsically indispensable. Because instrumental indispensability cannot justifiably ground ontological commitments to just *any* project, an instrumentally indispensable object x for an intrinsically indispensable project y cannot give us grounds to justify the pursuit of y . For example, a belief in the afterlife may be instrumentally indispensable for the project of explaining what happens to people's souls when they die, but this instrumentality does not give us reason to believe that the afterlife exists. And so Enoch claims the type of indispensability that applies to normative deliberation is that which requires us to make *ontological commitments* about the object of indispensability.³⁶

Enoch claims at least two intrinsically indispensable projects exist: the explanatory project and the deliberative project.³⁷ Enoch argues that we are explanatory beings that crave explanations.³⁸ When we observe the natural world, introspect, and go

³⁵ Ibid. 69. That is, the intrinsically indispensable project is impossible without its objects of instrumental indispensability.

³⁶ Ibid. 32. But one is not required to *believe* in the objects of indispensability.

³⁷ Ibid. 70.

³⁸ Ibid. 51.

about our day-to-day lives, we cannot help but seek out, offer, and construct explanations. We might be able to stop ourselves from engaging in the explanatory project temporarily, but Enoch maintains that our need to create and appropriate explanations is generally inescapable. It is an integral part of human experience.

Enoch believes the same is true for the normative deliberative project. When we observe the natural world, introspect, and go about our day-to-day lives, we deliberate constantly. We continually ask ourselves what we should do, what we should believe, how we should act, and how to reason. Should I go to war? Should I marry my partner? Should I vote for this politician? These are all normative, deliberative questions, and deliberating about these normative questions is also an integral part of human experience. We might be able to stop ourselves from engaging in the deliberative project temporarily, but Enoch argues that engaging in the deliberative project is also non-optional. Enoch's claim is this: if we engage in an intrinsically indispensable project, we are *justified* in believing that objects instrumental to that project exist.³⁹

For example, if the explanatory project is intrinsically indispensable, then the fact that the universe is explainable is instrumentally indispensable for it. Recall that instrumentally indispensable objects are those which cannot be eliminated without undermining our reason(s) for engaging in the project in the first place. If our reason(s) for engaging in the explanatory project are to explain the natural world, giving up on the fact that the universe is explainable *at all* renders the explanatory project impossible, or

³⁹ Ibid. 83. It is important to note here that the justification is both pragmatic (for actually engaging in deliberation) and epistemic (under the idea that our belief-forming methods are epistemically justified though not necessary reliable, if reliability here is connected with the truth conduciveness). Furthermore, *belief* in these normative facts is not required for Enoch's theory.

at least irrational. Thus we cannot give up on the fact that the universe is explainable, because doing so would undermine the explanatory project altogether.

Enoch argues we can say something similar about the normative deliberative project: if the normative deliberative project is intrinsically indispensable, then normative facts are instrumentally indispensable for it. If our reason(s) for engaging in the deliberative project are to make the *correct* normative decision, giving up on truth-makers for their correctness (i.e. normative facts) renders the project of deliberation impossible. That is, if we give up on normative facts, there is no correct decision to make, and the project of normative deliberation is undermined altogether.

For example, consider the normative decision of whether or not to have children. When we normatively deliberate about this issue, Enoch claims we assume that an answer - a correct answer, in fact - exists to the question, "Should I have children?" Normative facts determine whether or not having children is the correct choice, so we have an ontological commitment to the normative fact that we should or should not have children by engaging in normative deliberation. Deliberating about having children is simply impossible without a commitment to normative facts. We can make the decision to have or not have children arbitrarily, but then we would not be deliberating. We simply do not deliberate about an answer and dismiss it as arbitrary after the fact, since deliberation requires reflection and real consideration. I will defend this analysis in the next section.

3.2 Moral Deliberation Analysis

So far I have laid out the argument from indispensability. I have argued that normative facts are instrumentally indispensable for the intrinsically indispensable project of normative deliberation, and I have argued that belief in normative facts is justified since failing to utilize them in normative deliberation would undermine the project altogether. The next step then is to perform a thorough analysis of the psychology of normative deliberation by examining the goals and attitudes moral agents have in moral deliberation. In this section, I will show that moral agents utilize normative facts in normative deliberation. I will also offer support that normative deliberation is the intrinsically indispensable project I think it is. I will consider one type of normative deliberation, moral deliberation, and analyze it in what follows.

Consider the classic trolley problem. A trolley is hurtling down the tracks toward five workmen. It will most certainly kill them unless stopped or redirected. You stand by a lever that can divert the trolley onto a side path. However, there is one workman on this side path. If the lever is pulled, and the trolley is redirected onto the side path, it will most certainly kill the one workman. The moral question arises: do you pull the lever? Doing so will result in the death of one workman, while refraining to do so will result in the death of five. What should you do?

When engaging in moral deliberation, namely the cognitive treatment of moral questions, it is clear that we have certain goals in mind. Consider the moral decision of whether or not to pull the lever in the trolley problem. In the previous section, I argued denying normative facts would undermine the deliberative process entirely. This is because the goal of deliberation is to arrive at the correct answer to the normative

question we are trying to answer. When we ask ourselves whether or not to pull the lever, our goal is to arrive at the correct answer to this normative question.

How we arrive at that answer also shows that we rely on normative facts for normative deliberation. In normative deliberation, I argue that we arrive at the correct normative answer by *identifying* it or *discovering* it, and not by creating it.⁴⁰ Whenever someone engages in moral deliberation, they do not merely fabricate a framework for action on the spot. Consider again the trolley problem. When you ask yourself the normative question “Should I pull the lever?”, it is not as if you are constructing a framework for action right then and there. You ask yourself the question *to* find an answer. You weigh the different consequences your action will produce; you consider the value of human life and the distinction between killing and letting die, and you come to an answer that seems correct to you.

So when we engage in normative deliberation in order to decide what to do, we simply identify or discover a course of action that, to us, is intuitive to pursue. This realization provides for us the answer to our moral question. It tells us how we are supposed to act in the situation we are considering. The reason we decide to do what we do varies based on our intuitions and our beliefs about what we should do. Perhaps the best course of action produces the best consequences, or tracks our duty to our fellow man, or is the will of God, etc. Whatever intuitions we have, we operate under the assumption that they are correct until we reflectively scrutinize them.

That we scrutinize our intuitions regularly lends support to our ontological

⁴⁰ Ibid. 36. This is David Enoch’s point as well, and is a standard argument for normative facts.

commitment to normative facts as well. When confronted with evidence that runs counter to our intuitions, we do not merely change our intuitions - we *correct* them. This intuition business also permeates normative deliberation: we have certain intuitions about what we think we should do, and we assume that these intuitions track reality. Thus when we engage in normative deliberation, we ask ourselves normative questions, and we assume these questions have normative answers. The goal of normative deliberation then is to identify the correct answers to these normative questions. Not any old answer will do.

However, it is important to note that engaging in normative deliberation is not akin to mere picking, and Enoch elucidates this as well.⁴¹ For example, let's imagine that you go to the grocery store to pick out a cereal for breakfast. Picking out a cereal is not an intrinsically indispensable project - we can easily choose to pick out something else for breakfast, or nothing entirely. It may be that normative facts exist that ground our decision to pick out a cereal (i.e. the correct decision if we are having stomach problems is a cereal with high fiber), but our decision to pick is not undermined by denying normative facts. We are not required to take on any ontological commitments about cereal, in existential angst, when at the grocery store. When we engage in normative deliberation, we must assume normative facts exist. We mull the normative problem over in our mind, reflecting and considering it seriously. This is not required for picking out a cereal. Psychologically, we engage in normative deliberation differently than when we engage in mere picking.⁴²

⁴¹ Ibid. 37.

⁴² Ibid. 37. The duration and seriousness of the deliberation itself is not sufficient for characterizing moral deliberation, because we can also seriously deliberate about events at length that are not inherently moral (i.e. buying a car or what book to read). Rather the object of the deliberation (i.e. the moral problem) is what necessitates such serious and prolonged reflection in the first place.

In short I have shown that normative deliberation has illuminated certain facts about our psychologies: when we engage in normative deliberation, we assume correct answers exist to our normative questions. The fact that we trust our intuitions about what we should do and reflectively scrutinize them when challenged also lends support to our commitment to normative facts. Normative deliberation is not mere picking since it is apparent that we do not treat matters of normative deliberation as matters of mere preference: we seek to make the correct answer, and we do not treat this answer as arbitrary. How we feel and react to moral events also supports the view that we rely on normative facts for normative deliberation. I analyze this claim in the next section.

3.3 Moral Reactive Attitudes Analysis

In 1960 philosopher P. F. Strawson published an influential article titled “Freedom and Resentment” in which he sought to answer the following question: given the existence of determinism, is it rational to engage in moral responsibility ascription, praise and blame, and punishment and reward? Ultimately Strawson concluded that the question itself was nonsensical: the very *existence* of moral reactive attitudes shows that ascriptions of moral responsibility are actual and depend on no general metaphysical requirements.⁴³ For Strawson, morality (i.e. our holding each other responsible for moral actions) really consisted in our reactionary attitudes, so the question of rationalization conflicts with our actual moral commitments and moral behavior: “it is useless to ask whether it would not be rational for us to do what it is not in our nature to (be able to) do.”⁴⁴ Strawson claimed

⁴³ Strawson, “Freedom and Resentment,” 8.

⁴⁴ Ibid. 11.

that “reactive attitudes,” like gratitude, resentment, sympathy, anger, etc,⁴⁵ are naturally existing entities and functions of our social “transactions”⁴⁶ with other individuals in the moral community. Thus he concluded it is a fact that our reactive attitudes exist, and - as my undergraduate adviser succinctly put it - they are not up for grabs.

It is important to note that Strawson’s conclusion was built on two important premises: (1) moral responsibility exists and is grounded in our reactive attitudes; and (2) the existence of determinism *itself* in no way undermines the existence and cultivation of these reactive attitudes. More importantly, Strawson based these premises on one critical assumption: the phenomenological fact that we have certain moral attitudes is the vehicle which drives our justification for believing in moral responsibility.

This point is not far from Enoch’s conclusions, and so we can draw a parallel between Strawsonian reactive attitudes and Enochian normative facts. Both philosophers recognize some moral aspect of our psychologies is inherent in human nature. For Enoch, this is the non-optional project of normative deliberation. For Strawson, this is the non-optional possession of moral reactive attitudes. It is my claim that the existence of moral reactive attitudes also presupposes the existence of normative facts. When we have a reaction to a moral event, this reaction occurs when we evaluate the moral event against our intuitions about what is moral and immoral (i.e. against a particular class of normative facts). As we have seen, we assume this intuition tracks objective reality. So our moral reactive attitudes are functions of our intuitions which we assume track normative facts. But a question remains: in a deterministic universe, is it rational to hold

⁴⁵ Ibid. 4.

⁴⁶ Strawson, “Freedom and Resentment,” 3.

the reactive attitudes we do?

The argument from moral luck challenges this notion: if all that we do is determined, can we reasonably praise or blame someone for an action that is ultimately beyond their control? For example, suppose I steal a necklace from my friend Betty. Given determinism, it appears that my stealing the necklace is determined and outside of my control, a product of events in a causal chain to which I neither contributed nor for which I am responsible. Is it rational for any one person to blame me then for stealing Betty's necklace?

One response to this objection is to echo Strawson: it is clear that we have reactive attitudes of praise and blame. To consider whether or not having them is rational with the existence of determinism misses the mark. We do have them, and it would be impossible for us to rid ourselves of them. This response dismisses the possibility that moral responsibility is grounded in some metaphysical fact, sidestepping the objection. However, I find this response unsatisfactory, because I *do* want to ground moral responsibility in metaphysical fact.⁴⁷ Philosopher J. J. C. Smart offers another response in his article "Free Will, Praise, and Blame." Here Smart offers a revisionist account of praise and blame that is reconcilable with determinism.

Smart argues our reactive attitudes have both moral and non-moral functions, and that we employ them in three distinct ways: (1) to "grade"⁴⁸ individuals; (2) to encourage a specific type of action; and (3) to ascribe moral responsibility. Smart further distinguishes between praise and dispraise, which he argues have two unique non-moral

⁴⁷ I explicate my own views on the moral luck problem, which I call "moral chanciness" later in the essay.

⁴⁸ Smart, "Free Will, Praise and Blame," 303.

functions. The first non-moral function is gradation. When we praise or dispraise someone for something, we are merely commenting on the nature of the person, or how they are like. For example, suppose my friend Melissa is an expert pianist. When I praise her for her piano skills, Smart argues I am offering a positive grade toward her playing the piano. If my friend Melissa helps an old woman across the street, I am praising her in the same way by offering a positive grade towards her action of helping the old woman.

The second non-moral function is to encourage a specific class of actions. If I see my friend Melissa playing piano, I might praise her so that she will play piano more often in the future. Similarly, if I see my friend Melissa helping an old woman across the street, and I think this is a good thing, I might praise her so that she continues to help old women across the street in the future. However, Smart argues that blame has one unique moral function: to ascribe moral responsibility to a moral agent. When we blame someone for doing something, we are not merely grading their character; we are commenting on their moral status, and we are ascribing to them moral responsibility.

Smart's revisionist answer to the determinist objection is to offer a type of moral responsibility that is considerably weaker than other theories. When pressed, Smart concedes that moral responsibility is ultimately impossible in a deterministic universe. For he argues that when we praise, dispraise, and blame, we are not commenting on the metaphysical fact that the moral agent is responsible for the act she commits. While Smart goes too far in ultimately rejecting moral responsibility given determinism, there is merit to Smart's response. I do believe that his analysis of our moral and non-moral uses of praise, dispraise, and blame are correct and applicable in compatibilist theories.

However, I believe that we can still utilize these moral and non-moral uses in conjunction with our reactive attitudes to comment on true metaphysical moral responsibility.

The purpose of providing the philosophy of Strawson and Smart is twofold: to strengthen the case for our ontological commitment to normative facts, and to expose a larger problem in the contemporary moral-responsibility determinism debate. Smart denied metaphysical moral responsibility because he lacked a convincing compatibilist theory of free will and determinism. The worry is that being morally responsible for an action necessarily determines whether or not that person should be punished or rewarded for performing that action. While I agree the two are connected, they are entirely separate dimensions of philosophy. In the next section however, I will offer a theory of moral responsibility that takes advantage of this distinction, and I will show that my theory is strong and holds up to three incompatibilist objections by contemporary philosophers Saul Smilansky and Ishtiyaque Haji. I will argue that this compatibilist theory is not shallow compared to incompatibilist theories; nor should the compatibilist be charged with “complacent compliance” of accepting non-libertarian notions of justice and fairness. I also challenge one particular notion of the moral chanciness problem which arises from the consideration of possible worlds.

3.4 A Compatibilist Theory of Moral Responsibility

Recall again the discussion of normative deliberation: normative facts are instrumentally indispensable for the intrinsically indispensable project of normative deliberation. In the same way, since moral facts and deliberation is just a subset of the larger class of

normative facts and deliberation, I argue that *moral* facts are instrumentally indispensable for the intrinsically indispensable project of *moral* deliberation, the cognitive treatment of moral problems.

The compatibilist theory of moral responsibility I wish to elucidate is this: to determine whether or not an agent is morally responsible for an action is to compare said action with the moral facts we assume exist when we engage in moral deliberation.⁴⁹ Moral responsibility is then grounded in our ontological commitment to normative facts. My goal is not to explain the *origin* of these normative facts (i.e. whether or not they are constructivist in nature). Nor is my goal to elucidate what the normative facts *are* (i.e. what the normative facts tell us). It is simply to account for our ontological commitments to normative facts and our justification for believing in them by engaging in the non-optional project of moral deliberation.

Whether or not we may reasonably subject a moral agent to the appropriate moral reactive attitude(s) (i.e. praise or blame) is grounded in the adherence (or lack thereof) of the moral agent's voluntary actions to the normative facts.⁵⁰ Here voluntary actions are those which are functions of our own will. Voluntary actions may be influenced by other people and things, but ultimately we decide what course of action to take. That is, voluntary actions are not forced upon us. For example, if a gust of wind blew George

⁴⁹ Some may wonder if I am “giving up the game” so to speak completely to theories of moral psychology. My intuition is that, however, with proper and focused metaethics, these normative facts can be discovered and integrated neatly already into our pre-existing moral psychologies. So while our ontological commitment to normative facts depend on those objects with which we interact with using our psychologies, the facts themselves do not depend on our psychologies for their existence.

⁵⁰ However, we can subject any moral state of affairs to the appropriate reactive attitudes, even if they are not direct functions of any one particular person. We are within our reason to react morally to moral states of affairs if the desires of the moral community are to abide by moral states of affairs which embody goodness and love. Being morally disgusted with the holocaust for example is an instance of this.

over, who in turn knocked poor Margaret to the ground, this action would be involuntary. In this instance we could not reasonably subject George to blame. However, if Smith was standing next to George, urging him to knock poor Margaret to the ground, and George did so, this action would be voluntary. In this instance we could reasonably subject George to blame because it is ultimately George who decides which action he should take, even though he is influenced by Smith.

However, deciding whether or not to punish or reward a moral agent is determined separately from moral responsibility ascription and the employment of certain reactive attitudes. I am *not* arguing punishment and reward should follow necessarily from an agent being morally responsible for an action. Determining whether an agent is morally responsible for an action is entirely separate from determining whether or not an agent out to be punished or rewarded for performing said action. The existence of practical desert-undermining factors may preclude issuing punishment and reward on moral agents.⁵¹

My compatibilist theory of free action takes the form of the compatibilist theory explicated by philosophers John Martin Fischer and Mark Ravizza. The control necessary for free action, and thus moral responsibility, I believe is a certain kind of causal control, or proximal control, that Fischer and Ravizza call “guidance control.”⁵² According to

⁵¹ Practical desert-undermining factors constitute those things that have significant bearing on one’s moral actions and thus the moral community which the moral community has decided undermines the need for retributivism. If the purpose of retributivism (i.e. punishment and reward) is to discourage or encourage a specific class of actions, practical desert-undermining factors recognizes that implementing retributivism would be fruitless to the formation of one’s moral character or the the protection of the moral community.

⁵² Fischer, “Frankfurt-Type Examples and Semi-Compatibilism,” 307. For analyses and defenses of guidance control as it relates to moral responsibility, see Fischer (1994) and Fischer and Ravizza (1998). I do not offer a defense of guidance control here. My concern are broader arguments for compatibilism.

Fischer, guidance control requires that an action is generated by the moral agent, and the moral agent is reason responsive.⁵³

My compatibilist theory mirrors Fischer's "semi-compatibilism,"⁵⁴ which argues that in order to analyze and make sense of moral responsibility, we ought to analyze the actual sequence of moral events rather than the possible sequence of events afforded to us by *PAP*. Free action depends on the ability for moral agents to guide their moral actions in light of their intentions, desires, and reasons for acting. Whether or not these actions adhere to moral facts determines the charge of the moral action (i.e. whether they are moral or immoral). I believe this because I think the requirements for intuitive notions (i.e. libertarian) of free will and moral responsibility are impossible to achieve: self-origination and/or the capacity to determine one's actions in the absence of determinism.⁵⁵ I will not defend my compatibilist theory against guidance control-specific objections here.

However, there are several important broader objections to compatibilism. I will consider three objections to my compatibilist theory. The first two are summarized by philosopher Saul Smilansky, who argues that compatibilist theories fail to capture the true essence of moral responsibility. These objections concern the shallowness of

⁵³ Ibid. 307.

⁵⁴ Ibid. 307.

⁵⁵ Self-origination (or alternatively "ultimate origination") is discussed by philosophers Peter Van Inwagen and Saul Smilansky. I believe total self-origination in the libertarian sense is impossible. Similarly, I believe that libertarian notions of free will require that the moral agent has both the capacity to determine their own moral actions and some freedom from determinism. I believe these two conditions are irreconcilable: if indeterminism exists, we do not have the proper control necessary for moral responsibility. However, if causal determinism is false, we do not have the ability to determine our actions. It would seem then that both causal determinism must exist, and as agents we must be able to generate actions that are not subject to it. This is impossible on most libertarian views, including event-causal incompatibilist views. Agent-causal views might have the upper hand here, but I do not analyze them in this project.

compatibilist theories of moral responsibility, and the complacent compliance compatibilists engage in when ascribing moral responsibility and punishment and reward to moral agents. The third objection I will consider is advanced by philosopher Ishtiyaque Haji, who argues that it is a matter of moral chanciness as to which world (indeterministic or deterministic) into which a moral agent is born. I will take these arguments in turn.

Smilansky argues that any theories of moral responsibility have two requirements for moral responsibility ascription: (1) authorship of one's moral actions; and (2) authorship of one's self. Traditionally this second requirement⁵⁶ has come under fire. Many have argued that the second-order authorship requirement is impossible to fulfil. The argument is that, in order to be the author of one's self, there must exist a second layer of self generating this first part. However, in order to be the author of one's second layer of self, there must exist a third layer of self generating this part, and so on. Thus an infinite regress ensues, rendering the second-order author requirement unintelligible. In essence, "nothing can be *causa sui*."⁵⁷

For example, consider your own personality. When asked whether you *are* your personality, you might say that your personality constitutes you, but it is not entirely you. I think this is the intuitive response. The idea is that you, your self, has some control over your personality. You can make changes to it, and you can modify it. But what exactly are *you* then? This self is the second layer, generating your personality above it, but we believe (contradictorily it seems) that we can make changes to our self too. The question then is who or what is making these changes? The answer is that realistically there are

⁵⁶ Call this the "second-order authorship requirement" and the first requirement the "first-order authorship requirement."

⁵⁷ Strawson, "The Impossibility of Moral Responsibility," 5.

only so many layers to the self, and it has to stop somewhere.

For compatibilist theorists, simply dropping this second-order authorship requirement seems plausible. However, Smilansky believes that compatibilist theories become “shallow”¹⁴ if they do this. To do so, Smilansky argues, implies that the compatibilist is not the sole author of his own self - and he thereby grants that his personal identity is molded by people and things outside of him. Smilansky worries that, once we regard people and things outside of us as responsible for who we are, we cannot ascribe moral responsibility at all. For Smilansky, dropping the second-order authorship requirement entails dropping the first-order authorship requirement as well, because the cause of a moral agent’s action can always be traced back to something earlier in the causal chain. The moral agent is then just another link in this causal chain and nothing more.

However, I disagree that dropping the second-order authorship requirement entails also dropping the first-order authorship requirement. Though the self may be ultimately determined by forces outside of us, the self still engages in moral decision making. The self being a link in the causal chain does not deprive the self of its own capacity to determine voluntarily future events in the causal chain. In my theory, a sufficient explanation for why Greg pushed Margaret to the ground consists in Greg guiding his own actions, intentionally and voluntarily exercising his will. We do not need to reference anything outside of Greg to explain his actions, because the decisions he makes can still be necessary and sufficient for the outcome we are considering. Compatibilist theories do not fail to make sense of “true” moral responsibility, because

we need not always regard morally responsible agents as deserving of punishment or reward for their actions. Dropping the second-order authorship requirement does not absolve moral agents from making moral decisions.

Smilansky also argues that compatibilist proponents cannot live up to libertarian-based notions, and so what compatibilist proponents must offer are “shallower sorts of meaning” for justification and fairness.⁵⁸ However, this conclusion is misleading. On the one hand, it is not clear that this is a problem for the compatibilist at all. We are only committed to the belief that shallower sorts of meaning are problematic if we assume libertarian intuitions about justification and fairness are correct. Smilansky offers no reasons to accept these intuitions. On the other hand, the argument fails against my compatibilist theory. The very existence of normative facts rebuts this objection. I claim there is no stronger, more forceful meaning for notions of justice and fairness than those which track normative facts. On my view, compatibilism offers its proponents an ultimate source of justification and fairness. Thus I do not believe my theory of compatibilist falls victim to the shallowness objection.

The second objection Smilansky raises is the argument from “complacent compliance,”⁵⁹ which attacks the compatibilist on two fronts: (1) by claiming compatibilist theories permit unjust punishment; and (2) by administering this unjust punishment anyway. For example, consider again the above case in which I steal my friend Betty’s necklace. Smilansky would argue that compatibilism is weak on two counts: first, it acknowledges that I am not metaphysically responsible for stealing the

⁵⁸ Smilansky, “Free Will, Fundamental Dualism, and the Centrality of Illusion,” 493.

⁵⁹ Ibid. 493.

necklace (my theft being causally determined by forces outside of my control), and second because the theory seeks to punish me anyway. To Smilansky, this just seems intuitively wrong.

However, Smilansky's generalization causes this argument to be a non-starter against my theory. My compatibilist theory takes no prisoners: ascription of moral responsibility is utilized, justified, and intuitive in my compatibilist theory. The existence of normative facts and our ontological commitment to them entails and justifies the existence of the genuine metaphysical moral responsibility that Smilansky is after.⁶⁰ Second, justification of punishment inevitably follows from the first point. Surely moral agents deserve to be punished if they utilize normative facts in moral deliberation and exercise their will maliciously and voluntarily. If this response is unsatisfactorily *ad hoc*, I throw my towel also in the ring of compatibilists who justify punishment based on practical consequences, but that is a topic for another paper.

In fact, Smilansky's objections to compatibilism are weakened by his illusionist account of moral responsibility, which is the view that moral responsibility does not truly exist, but the illusion that exists is necessary and practical. As a proponent of this view, Smilansky is guilty of having his accountability cake and eating it too. Smilansky cannot find the compatibilist morally responsible for complacent compliance in good faith if he does not believe it exists. Similarly, he cannot earnestly find compatibilist theories of

⁶⁰ Whether or not this moral responsibility is actually *attainable* is another question entirely. Of course I believe that our conception of moral responsibility ought to be revised, its current definition being a function of mistaken intuitions about possible self-origination. Under my compatibilist theory, if we do revise our theory of moral responsibility to be a function of guidance control, then moral responsibility both exists and is attainable. I do not offer an argument in this honors project for revising our current theory of moral responsibility, nor do I offer an argument for moral responsibility inevitably following the existence of moral facts. I take this as intuitively given.

moral responsibility shallow if he does not believe that moral responsibility exists. The argument from Smilansky thus fails on more than one count, and my theory remains strong and intuitive.

The objection from Ishtiyaque Haji is of a slightly different flavor. Haji argues that compatibilist theories also suffer from the problem of moral chanciness. Haji's argument is twofold: (1) given that all events are determined in a deterministic universe, compatibilist accounts of free will also suffer from moral chanciness; and (2) if moral agents lack power or control over the sort of world they are born into (i.e. deterministic or indeterministic), the world into which a moral agent is born cannot affect the degree to which they are morally responsible.⁶¹ I will examine both of these objections in turn.

Haji's first objection relies on the idea that true accounts of moral responsibility require either complete authorship of an agent's actions (i.e. self-origination), or complete control over an agent's actions. On a deterministic worldview, self-origination is impossible, because a moral agent's self is inevitably a function of their upbringing, biology, and other environmental factors. Complete control over one's actions is also impossible, because the type of control necessary for moral responsibility is that which is not affected or produced by causal determinism. Because all events in the causal chain are determined, the only control a compatibilist theory can offer is a sort of control that is compatible with causal determinism - not outside of it.

However, philosopher Galen Strawson has shown quite convincingly that the self-origination requirement for moral responsibility is impossible to fulfil, because it

⁶¹ Haji, "Alternative Possibilities, Luck, and Moral Responsibility," 270.

leads to an infinite regress.⁶² If we are forced to abandon moral responsibility or this requirement for moral responsibility, I opt to abandon the requirement. Furthermore, I believe that our current theories about moral responsibility are influenced by our mistaken intuitions regarding self-origination and ultimate control. I have proposed a revised theory of moral responsibility, which makes a distinction between moral responsibility and moral desert, and does not fall prey to these types of libertarian objections concerning being the ultimate source of one's actions or having ultimate control of one's actions. If this revised theory is correct, Haji's first objection has little bite.

In any case, it is important to note that this first objection from moral chanciness does not solely damage compatibilist theories, but as I have shown in chapter 2 it damages incompatibilist theories as well. Any philosopher involved in the free will and moral responsibility debate then has stake in rebutting the moral chanciness objection; the burden does not fall merely to the compatibilist. It is a problem that plagues all moral theories to my knowledge. Thus while I do not have a solution to the problem of moral chanciness, I pass off the torch to more clever, astute minds.

Haji's second objection concerning moral chanciness is a little stranger. Haji explicates a principle by the contemporary philosopher Michael Zimmerman, which states that our degree of moral responsibility cannot be affected by things outside of our control. Call this the Principle of No Control:

PNC: The degree to which we are morally responsible cannot be affected by what is not in our control.⁶³

⁶² See the previous citation on Strawson 57.

⁶³ Zimmerman, "Taking Luck Seriously," 3.

PNC is meant to capture the intuition that, if one is powerless or lacks control over something *a*, and one is morally responsible for some other thing *b*, then *a* cannot affect the degree with which one is morally responsible for *b*.

For example, if Jonathan volunteers his time at a homeless shelter, feeding and educating the poor, he is morally responsible for making sure that the homeless receive meals during his shift. However, while Jonathan is not volunteering, he is not morally responsible for providing for the homeless. Thus insofar as Tim, Jonathan's coworker, hordes the food for himself and refrains from feeding the homeless, Jonathan cannot be held responsible for the deteriorating conditions of those in the shelter while he is not working. The degree to which Jonathan is morally responsible for feeding the homeless cannot be affected by Tim's actions.

Haji notes that it is not up to any moral agents what kind of world into which they are born. It might have been the case that they were born into an indeterministic world, or it might have been the case they they were born into a deterministic world. Because of this fact, compatibilist theories also have a problem with moral chanciness: had it been the case that agents were born into the sort of world that renders them morally responsible for their actions, this is just dumb luck because these moral agents could not have control over what universe they are born into, and *PNC* tells us that the degree to which they are morally responsible cannot be affected by considerations like these.

I am not convinced that *PNC* is correct.⁶⁴ However, what I want to contest is

⁶⁴ Zimmerman argues that moral agents (i.e. George and Georg, let's say) can be equally morally responsible for actions (i.e. attempting to kill Henry and Henrik, respectively, let's say) regardless of whether or not Henry or Henrik is actually killed, perhaps due to situations involving circumstantial or constitutive luck. Zimmerman traces moral responsibility to the counterfactual "if a moral agent *S* would have freely done a moral action *a*, then *S* is morally responsible for *a*." This leads Zimmerman to conclude

Haji's claim that it is a matter of moral chanciness concerning the kind of universe into which someone is born. It is not apparent to me that being born into a deterministic or indeterministic universe is a matter of moral chanciness at all. Here I am not claiming that moral agents have control with which to determine their universal origins, but I am contesting the fact that lack of control *per se* constitutes moral chanciness.

It is my intuition that, to be morally chancy, one must be morally chancy with respect to moral actions, not necessarily moral states of affairs. If we assume that being born into an indeterministic or deterministic is a moral state of affairs,⁶⁵ it does not follow that this also constitutes moral action. In fact, I would argue that one's birth is not a moral action that moral agent performs at all. It is a misuse of the phrase 'moral chanciness' to equate the chanciness of a moral event with the chanciness of a moral action. It is the latter that concerns compatibilist theories, not the former. If a moral agent is responsible for a moral state of affairs, it must be because that state of affairs is the result of some sort of moral action that moral agent performed. And certainly the moral state of affairs of the kind of universe into which a moral agent is born is not a state of

that issues of moral luck are irrelevant to moral responsibility. However, I find this treatment of moral responsibility questionable. Part of what makes a person morally responsible for their actions is not only what the intentions of those actions are, but also the consequences of those actions, which most of the time are under the control of the moral agent. If it is a fact that killing someone is morally wrong, then George is obviously more responsible than Georg, because he committed a moral action that Georg did not: killing Henry. We can say that both agents are equally morally responsible for *attempting* to kill Henry and Henrik, but we cannot say that both agents are equally morally responsible *in toto*. I am not convinced that Zimmerman's treatment of moral responsibility is correct, given that he also thinks agents can be morally responsible without having done any action whatsoever. It is my intuition that Zimmerman incorrectly categorizes the relation between moral responsibility and what moral agents are responsible *for*. It is not clear to me that Georg is morally responsible for everything that George is morally responsible for, and thus I shall reject *PNC* at this present time.

⁶⁵ That is, a state of affairs concerning, related to, or connected with moral responsibility - a state of affairs with a moral charge (i.e. a state of affairs substantiated by good and evil, adhering to or breaking moral law). A moral state of affairs contains a moral charge, but this moral charge does not reference one particular person or action - it references the macroscopic or totality of good or bad moral events occurring.

affairs that is resultant of any moral action performed by that agent.

Furthermore, I am not convinced that the kind of universe into which a moral agent is born is a moral state of affairs at all. It is not clear to me that a moral state of affairs concerning only the fact of whether the universe is indeterministic or deterministic has any moral charge whatsoever. There is nothing inherently moral or immoral about this characteristic of the universe, so the kind of universe into which one is born does not constitute a state of affairs for which any (human) moral agent could be responsible or morally chancy.

Haji argues that the kind of universe into which a moral agent is born however undermines the responsibility an agent could have in that universe, but I disagree. In the Charles and Charles* examples, indeterminism seems to be a responsibility-undermining factor because it undermines the responsibility that Charles and Charles* could have had in their respective universes. However, for a moral agent being born into one of these universes, their responsibility is in no way undermined in the same way.

I disagree with Haji in thinking that the kind of universe into which a moral agent is born substantiates a responsibility-undermining factor for that moral agent. Either the universe a moral agent is born into precludes moral responsibility or allows it. If the moral agent is born into the universe that precludes moral responsibility, there is no status of moral responsibility in that universe for which he can be chancy. If the moral agent is born into the universe that allows for moral responsibility, he still cannot be chancy because it was impossible for him to have been born in any other universe. One can only be chancy with respect to events that could have been otherwise. One's birth is not an

event like this.⁶⁶

3.5 Conclusion

In short I have argued that moral deliberation and participation in moral reactive attitudes is inescapable. I have argued that normative facts are instrumentally indispensable for this project of moral deliberation, and as such we are justified in believing in their existence. I have offered a revisionist account of our notions of praise and blame, partly in the spirit of J. J. C. Smart. Using these ideas, I have set out a compatibilist theory of moral responsibility that grounds moral accountability in these normative facts. Reasonable operation of the appropriate moral reactive attitudes in response to an agent's moral action must reference the agent's adherence (or lack thereof) to the moral facts. Our justification for having strong reactive attitudes and ascribing moral responsibility are grounded in our ontological commitment to moral realism, and our capacity for free action consists in the compatibilist account of guidance control given by John Martin Fischer. I have argued that my compatibilist theory stands up to three common incompatibilist objections: the argument from shallowness, the argument from complacent compliance, and the argument from moral chanciness from possible worlds.

⁶⁶ To elaborate, consider Yvonne and Yvonne* born into worlds w_1 and w_2 , which are deterministic and indeterministic respectively. To consider whether or not Yvonne and Yvonne* could be morally chancy, we must make reference to the actions of Yvonne and Yvonne*, the extent of their control, and their personal identities. Any appraisal of a moral agent necessary is a function of some aspect of their personal identity. We cannot make sense of the moral chanciness of Yvonne and Yvonne* *if they do not exist*. But Yvonne and Yvonne*'s location in spacetime, including the specific worlds into which they are born, figure necessarily into their personal identity. They wouldn't be Yvonne and Yvonne* if they hadn't been born into worlds w_1 and w_2 , respectively. Thus it does not make sense to talk about the possible worlds into which Yvonne and Yvonne* could have been born, because we're not talking about Yvonne and Yvonne* anymore: we're talking only about our *ideas* of Yvonne and Yvonne*. So Yvonne and Yvonne* could not have been born into any other world than the ones into which they were born, and they cannot be morally chancy with respect to this event.

I understand that my account may not be totally convincing or satisfactory for proponents of moral subjectivism, or those who are not keen on accepting moral realism. Thus in the next section I propose two alternative, non-libertarian views of moral responsibility - Saul Smilansky's illusionism and Derk Pereboom's hard incompatibilism, influenced by Michael McKenna's conversational theory of moral responsibility. However, I show that these views also have problems, and I will ultimately conclude that they fail to offer us notions of moral responsibility that more tenable than my compatibilist theory.

Chapter 4: Alternative Views - Illusionism and Conversational Theory

4.1 Introduction

In this section I will consider and analyze two alternative views of moral responsibility which are incompatibilist but non-libertarian. These views are illusionism and conversational moral responsibility, by philosophers Saul Smilansky and Derk Pereboom, respectively. I will argue the illusionism of Smilansky is unattractive because it prescribes secrecy about the conditions of moral responsibility, and I will argue that the conversational theory of Pereboom is unattractive because it requires the existence of a plurality of moral agents with finely-tuned reactive attitudes to generate meaningful moral ascriptions.

4.2 Smilansky's Illusionism

Philosopher Saul Smilansky considers the problem of free will and moral responsibility in his article “Free Will, Fundamental Dualism, and Centrality of Illusion.” Smilansky begins his discussion of free will by stating three questions that concern the moral responsibility-determinism debate. He says, “I believe that the best way to understand the problem of free will is as a conjunction of three questions: ... the libertarian Coherence or Existence Question ... the Compatibility Question ... [and] what are the consequences of the undoing of both libertarianism and (in part) compatibilism?”⁶⁷ The first question reduces to: ‘Is there libertarian free will?’ and the second reduces to: ‘Is moral responsibility compatible with the absence of libertarian free will?’

Smilansky offers a pessimistic answer to the libertarian Coherence Existence Question. Without going into a deep analysis of the free will debate, Smilansky argues that “the conditions required by ... libertarian free will ... are self-contradictory and hence cannot be met.”⁶⁸ Smilansky believes that the control required for moral responsibility is not generated by indeterminism present in libertarian theories.⁶⁹ Similarly, he regards that the control requirement itself is impossible to meet, since the self-determination of one’s moral character⁷⁰ produces an infinite regress.

Smilansky then considers the Compatibility Question, analyzing compatibilist and hard determinist answers to the problem of moral responsibility and determinism in turn. Smilansky argues that compatibilism fails on two counts. First, it offers a shallower sort of meaning and justification for justice and moral responsibility than its libertarian

⁶⁷ Smilansky, “Free Will, Fundamental Dualism, and the Centrality of Illusion,” 490.

⁶⁸ Ibid. 491.

⁶⁹ My analysis of event-causal incompatibilism supports this view as well.

⁷⁰ Sometimes known as “ultimate origination” in the literature. I use this phrase as well in the essay.

counterparts, which do not match up with our intuitions.⁷¹ Second, the compatibilist remains “complacently compliant” with failing to recognize this injustice of lack of fairness and desert.⁷² That is, Smilansky charges the proponent of compatibilism with being untroubled about punishing and ascribing moral responsibility to moral agents for actions outside of their ultimate control.

Smilansky continues his analysis by arguing that hard determinism - the denial of moral responsibility - also fails. He reasons that there are certain instances where a compatibilist basis for moral responsibility is plausible. However, hard determinism opposes all such cases. Smilansky argues that “we want to be members of a Community of Responsibility where our choices will determine the moral attitude we receive, with the accompanying possibility of being morally excused when our actions are not within our reflective [proximal] control.”⁷³ Because hard determinism precludes this possibility and fails to generate convincing reasons to ascribe moral responsibility to agents at all, hard determinism then fails as a theory about moral responsibility in practice because it cannot account for our desire and our need to be morally responsible agents.

Finally, Smilansky considers a joint perspective that combines compatibilist and hard determinist theory, which he labels Fundamental Dualism.⁷⁴ Fundamental Dualism is a view that embraces the strengths of both compatibilist and hard determinist theories, arguing that both theories have useful contributions to make in both the moral responsibility-determinism debate and the social institution of moral responsibility

⁷¹ Ibid. 493.

⁷² Ibid. 493.

⁷³ Ibid. 495.

⁷⁴ Ibid. 495.

ascription and punishment. This view directly opposes the Assumption of Monism, which argues that philosophers in the moral responsibility-determinism debate must either affirm compatibilism or incompatibilism.⁷⁵ Smilansky ultimately rejects the Assumption of Monism in favor for Fundamental Dualism, citing that “... there is no conceptual basis whatsoever for thinking that the Assumption of Monism is necessary. Compatibilism and incompatibilism are indeed logically inconsistent, but it is possible to hold a mixed, intermediate position that is not fully consistent with either.”⁷⁶

This leads Smilansky to develop a theory of Illusionism about free will, moral responsibility, and determinism. Illusionism is the position that libertarian free will does not exist, but illusory beliefs that it exists are in place in society and play a largely positive role in moral responsibility ascription and punishment. Indeed these illusory beliefs, Smilansky argues, play an necessary role in ascribing moral responsibility and understanding praise, blame, and moral worth. While ultimate moral responsibility does not exist, it is imperative that we propagate the illusion of free will and moral responsibility to sustain our social and moral relationships within what Smilansky calls “the moral community.”

Smilansky argues illusionism is crucial for several reasons. First, it is the only non-libertarian theory of moral responsibility that allows for questions of innocence, and questions of innocence are required for a functional moral society. Smilansky writes, “if a moral system ... is to function well, [innocence] should be prevalent, almost instinctive.”

⁷⁷ However, compatibilist and hard determinist theories leave no room for questions of

⁷⁵ Ibid. 491.

⁷⁶ Ibid. 491.

⁷⁷ Ibid. 498.

innocence, and so they engage in patently unjust forms of moral treatment by punishing and rewarding moral agents for actions for which they have no control. Similarly, Smilansky doubts that moral communities without questions of innocence could function at all. So compatibilist and hard determinist moral communities will necessarily fall into chaos and disorder.

However, Smilansky is too quick with this objection. He offers no empirical evidence to support that moral communities based on compatibilist and hard determinist principles will fail. Similarly, my compatibilist theory argues that true moral responsibility and proper moral treatment follows naturally from normative facts and a semi-compatibilist notion of free will. It does not concede that moral agents are all equally guilty or innocent, and it rejects the libertarian requirements for moral responsibility, which Smilansky supports. Rather it is illusionism that regards all moral agents as equally guilty or innocent - it just refrains from explicating this fact. If true moral responsibility is impossible, and if innocence is then an illusion, illusionism's tendency to engage in differential treatment for moral agents is patently unfair or unjust in the same way Smilansky believes compatibilist or hard determinist moral treatment is. Thus illusionism falls victim to Smilansky's own criticism.

Similarly, Smilansky believes that only illusionism can foster a moral community in which the agents find themselves accepting of practices that will positively affect their future behavior. He believes that, for any cogent theory of moral responsibility to work, moral agents must hold themselves as responsible for their moral actions, but this runs directly counter to compatibilist intuition. If agents internalize the fact that they are not

ultimately morally responsible for anything, they will not take themselves and their moral actions seriously. Thus illusionism is the only theory that can generate this kind of attitude for moral agents

However, this objection is again too quick. My compatibilist theory does offer ultimate moral responsibility. It provides the most objective standard with which to govern one's actions, which depends on the existence of objective, universal, normative fact. Furthermore, even if a moral agent does internalize the truth of ultimate non-responsibility, one can still reflect about moral principles and use them to *practically* guide their actions. There is nothing about compatibilism that precludes this type of reasons responsiveness - that is, the capacity to reflect and act morally in light of the the social consequences for one's actions. In fact, compatibilist theories are built on this type of reason responsiveness, so they do not necessarily diminish the effects one feels from moral actions.

Smilansky also believes that illusionism is the only non-libertarian theory that can justify the existence of moral worth and moral value. He says, "from the ultimate hard determinist perspective, all people ... are morally equal ... there cannot be any means of generating a 'real' moral value."⁷⁸ Smilansky insists that internalizing this hard determinist perspective leaves moral achievements valueless, and true appreciation of moral actions is impossible if we recognize that the moral actions of all moral agents are determined. Compatibilism is unsatisfactory as well, because it offers little protection against these hard determinist intuitions. In short only illusionism can positively affect

⁷⁸ Ibid. 499.

the “very meaning we can find in our lives.”⁷⁹

Yet illusionism fares no better in this regard, since it fails to generate any ‘real’ value or worth as well. All illusionism is able to propagate is the *facade* of ‘real’ moral value or worth. If true moral value and moral worth depend on free will, illusionism cannot offer any more ‘real’ appraisals of moral value and moral worth than other non-libertarian theories. Furthermore, it is not obvious at all that value and worth in general depend on free will. Just as value or worth can come from natural abilities and characteristics (i.e. such as athleticism or artistry), so too can we place real moral value or worth on moral characteristics, regardless of whether or not they come about by free will (i.e. such as honesty, goodness, or kindness). In this way, my compatibilist theory can still generate and retain real moral value or worth.

Finally, Smilansky argues that internalizing the hard determinist perspective precludes moral agents from feeling compunction. Only illusionism can allow for this sort of feeling, as well as the existence of remorse and integrity. Smilansky argues that genuine feelings of moral responsibility are crucial to operating and regarding ourselves as morally responsible. Being a hard determinist causes a reduction in seeing not only others as moral agents, but seeing one’s self as a moral agent as well. In this sense, “feelings of remorse are inherently tied to the person’s self-perception as a morally responsible agent,”⁸⁰ and non-libertarian theories cannot rationalize or actualize this self-perception.

Here illusionism may have the upper hand against its compatibilist rivals.

⁷⁹ Ibid. 499.

⁸⁰ Ibid. 500.

However, we can still conceive of compunction arising from a compatibilist theory, or in situations in life in which the moral agent could not have done otherwise. For example, imagine a drug addict that relapses for the third or fourth time. Even though the drug addict is determined to relapse due to their desire for drugs, knowing that they are slaves to their desires does not lessen the guilt they feel for disappointing their friends and families again by relapsing. I am inclined to argue that guilt does not arise from knowledge concerning one's metaphysical circumstances (i.e. lacking free will), but rather the effects that their actions have on themselves and on the moral community. This point of compunction then remains inconclusive.

I have one more general objection against illusionism. Recall the tenets of illusionism: (1) the general population holds false beliefs about the existence of free will, moral responsibility, and determinism; (2) these false beliefs are vital to ascriptions of moral responsibility and participation in the moral community; (3) those who know the truth about moral responsibility and determinism (i.e. that determinism exists and free will and moral responsibility do not exist) should keep it to themselves.

Illusionism implies that false beliefs about ascriptions of moral responsibility are vital to participation in the moral community, and it prescribes that those who know the truth about moral responsibility, free will, and determinism should live and act in a way that is contradictory to our beliefs: that we act as morally responsible agents with knowledge that we are not ultimately responsible for anything. There are practical reasons for doing this, but it seems that any metaphysical theory that argues we ought to act in

ways that are contradictory to our beliefs lose serious philosophy brownie points.⁸¹ For this reason *I* find illusionism unsatisfactory.

For the above reasons I do not believe that Smilansky's illusionism is a viable alternative to compatibilism. If this is true, we ought not adopt illusionism as a theory of moral responsibility. Not only does it prescribe lying to lay people about the truth of their situation, but it also seems to prescribe acting in ways that are contradictory to our beliefs (those of us in the know, at least). Such a radical, nonstandard view of moral responsibility then should be discarded in favor of my more tenable compatibilist theory.

4.3 McKenna/Pereboom's Conversational Theory

Philosopher Derk Pereboom also considers a nonstandard view of moral responsibility and free will in his book *Free Will, Agency, and the Meaning of Life*. Pereboom describes himself as a hard incompatibilist. *Hard incompatibilism* is the view that neither free will nor moral responsibility is compatible with the truth of causal determinism.⁸² However, Pereboom espouses an optimistic view of life and moral responsibility without free will in reference to concepts of agency, morality, and the meaning of life. In fact, Pereboom argues that the hard incompatibilism is compatible with rational deliberation, practical moral responsibility ascription, and a system with dealing with criminal behavior and punishment.⁸³ It is this hard incompatibilist view that I will analyze here.

⁸¹ Or 'plausibility points,' according to David Enoch. See the closing chapters of his book *Taking Morality Seriously*. It is merely an intuition that drives this assertion. If a philosophical theory prescribes actions in real life that contradict its tenets, the theory itself seems hypocritical and questionable to say the least.

⁸² Pereboom, *Free Will, Agency, and the Meaning of Life*, 4. Pereboom himself remains agnostic about the truth of causal determinism, but he rejects the possibility for libertarian free will regardless.

⁸³ *Ibid.* 4.

Pereboom argues that the notion of moral responsibility at issue in the contemporary free will and moral responsibility debate is that of basic desert. Moral responsibility in the basic desert sense is the notion that a moral agent is morally responsible for an action if “the agent would deserve to be blamed or praised just because she has performed the action, given an understanding of its moral status.”⁸⁴ This is the traditional view of moral responsibility - requiring an incompatibilist, libertarian notion of free will - and it is at odds with contemporary notions of compatibilist and hard determinist accounts of free will and determinism

Broadly, incompatibilists argue that this basic desert notion of moral responsibility is undermined by the truth of causal determinism. They argue that a moral agent is only deserving of blame if they satisfy the requirements for moral responsibility (i.e. alternative possibilities, self-origination, or being the proper source of one’s actions), and causal determinism precludes one or more of these. Compatibilists in general question which of these are necessary for moral responsibility, and they reject the notion that determinism precludes the one or more of these requirement(s) for moral responsibility. However, Pereboom argues that this basic desert is not the only notion of moral responsibility that we should consider in moral responsibility ascription.

Pereboom maintains that there are various notions of moral responsibility that are not undermined by compatibilism and hard determinism and have not yet been considered seriously in the free will and moral responsibility debate. Pereboom analyzes one of these notions of moral responsibility, centered around the concept of blame, and

⁸⁴ Ibid. 127.

he argues that this notion of moral responsibility can be reconciled with the nonexistence of free will and the truth of causal determinism. Because most definitions of desert and moral responsibility involve the notion of basic desert, Pereboom offers a revisionist account of moral responsibility.⁸⁵

Pereboom proposes that grounds for moral blame and moral responsibility ascription should consist not in our notion of basic desert, but in three moral considerations: protection of potential victims, reconciliation to personal relationships and relationships with the whole moral community, and moral formation.⁸⁶ Pereboom argues that these moral considerations arise from the harmful nature of immoral actions. Because immoral actions have the potential to harm relationships and moral agents, we have stake in positively affecting moral relationships and moral agents within the moral community by fostering the formation of a good moral character, undoing harm to moral relationships, and protecting ourselves from harmful actions.⁸⁷

Following philosopher Michael McKenna, Pereboom proposes a *conversational theory of moral responsibility*, where moral action, responsibility ascription, and punishment is analogous to an ordinary conversational exchange between speakers of a natural language.⁸⁸ This view rests on the interdependence thesis, which asserts that facts about moral responsibility ascription depend in no way upon facts about agents being morally responsible.⁸⁹ Under this conversational model, the purpose of moral responsibility ascription is to communicate a moral agent's response to the quality of will

⁸⁵ Ibid. 128.

⁸⁶ Ibid. 134.

⁸⁷ Ibid. 134.

⁸⁸ Ibid. 132.

⁸⁹ I will comment on the interdependence thesis later in this section.

of another moral agent. The actions of a moral agent are indicative of this quality of will with which moral agent acts, and the quality of will is a function of the agent's desires, intentions, thoughts, beliefs, etc. For example, if Olivia steals my wallet, under this conversational model, the purpose behind my blaming her for doing so is to communicate my response to the moral action she just committed. By blaming Olivia for stealing my wallet, I am commenting on Olivia's quality of will, or her moral character. This kind of analysis is similar to the revisionary account of moral responsibility of J. J. C. Smart which I analyzed above.

In the words of McKenna, modified by Pereboom, the conversational model can be broken up into different stages: moral contribution, moral address, and moral account. The first stage is moral contribution, in which a moral agent performs an action that she implicitly conceives as being morally charged (i.e. good, bad, or indifferent) and indicative of a certain quality of will. The second stage is moral address, in which the moral agent is recognized by another moral agent as having performed an action. The third stage is moral account, in which a moral agent morally ascribes the agent that had performed the moral action. In this stage the agent having performed the action extends the conversation by offering an excuse or justification for their action, or an apology for the action altogether, and the other moral agent may respond by administering further punishment if the excuse, justification, and/or apology is lacking.

For example, imagine that Tom and Lucy are at a concert together seeing their favorite band. Between songs, Lucy leaves her seat to go buy a beer. When she returns, she gets bumped in the back and accidentally spills her beer all over Tom. In exasperation

Tom throws his arms up in the air and looks at her angrily. He blames Lucy for spilling the beer on him, and Lucy responds by offering an excuse: someone bumped her from behind. Tom mulls over this excuse and concludes that it is viable, and so Lucy is not deserving of blame. Tom brushes himself off and refrains from pursuing further punishment for Lucy's actions.

In the Tom and Lucy example, the purpose of the moral responsibility exchange is three-fold: protection of potential victims, reconciliation to personal relationships and relationships with the whole moral community, and moral formation. By addressing and blaming Lucy for her actions, Tom is exercising his right to call out Lucy's behavior for two reasons. He attempts to show Lucy that what she did was wrong, so that she can modify her behavior in the future by engaging in positive moral formation, and he attempts to safeguard himself and other moral agents from future harm by Lucy's actions. By engaging in the moral responsibility exchange together, Tom and Lucy attempt to reconcile their personal relationship with the moral action that Lucy has just committed.

Pereboom believes that this theory of blame can be extended to theories about praise as well. Consider Tom and Lucy again. However, imagine that this time when Lucy gets back to her seat, she sees that Tom is in a heated argument with another concert-goer. The concert-goer is attempting to steal Lucy's seat for himself, but Tom fends him off. When Lucy returns, she praises Tom for doing the morally right thing, for securing her seat which she paid for in her absence. In this case the moral exchange goes about just like before, except Tom is commended for his actions. This strengthens the relationship between Tom and Lucy and encourages positive moral formation in Tom's

character, thereby securing protection for moral agents in the future that might engage with Tom.

Pereboom defends this conversational theory of moral responsibility by considering the objection that blaming someone requires alternative possibilities, which is impossible in a deterministic universe. In response Pereboom distinguishes between two different types of ‘ought-statements:’ statements of axiological evaluation and specific action demands.⁹⁰ The former are statements that describe how a moral agent ‘ought-to-be,’ while the latter are statements that describe what a moral agent ‘ought-to-do.’⁹¹ According to philosopher James Hobbs, if we are making an axiological evaluation about a moral agent or a state of affairs (i.e. ‘if X ought-to-be the case’), we are saying that the moral agent has a reason to act in X-ing. This evaluation does not require alternative possibilities, so it does not rely on the assumption that ‘ought-implies-can.’ However, if we are making a specific action demand about a moral agent (i.e. that they ‘ought-to-X’), this demand does require alternative possibilities, so it does depend on the assumption that ‘ought-implies-can.’⁹²

Pereboom argues that the conversational theory of moral responsibility utilizes axiological evaluations about a moral agent rather than specific action demands. This is similar to the analysis of praise and blame of J. J. C. Smart. Pereboom calls this type of moral responsibility ascription axiological recommendation, where the ‘ought’ in this recommendation does not imply obligation to perform or refrain from certain actions.⁹³

⁹⁰ Ibid. 139.

⁹¹ Ibid. 139.

⁹² Ibid. 139.

⁹³ Ibid. 141.

This type of analysis requires that moral agents be reasons responsive so that they might modify their future actions or understand their immoral actions presently through conversational theory of moral responsibility.

Pereboom's proposal for moral responsibility then differs from Smilansky's illusionism in a variety of ways. Pereboom offers a revisionist account of moral responsibility that is forward-looking and influential in practice. For Pereboom, moral responsibility would then exist, but it would not be the kind of moral responsibility that we are used to. Instead of our theory of moral responsibility relying on our mistaken notion of basic desert, which is irreconcilable with hard determinism, it relies on the aforementioned moral considerations. Contra Smilansky, we can have a real theory of moral responsibility rather than a fake or illusory theory of moral responsibility, albeit a different theory.

Instead of attempting to cover up the truth about free will and moral responsibility, Pereboom's forward-looking theory prescribes that we attend to our false beliefs head on in attempts to *revise* our current theories of moral responsibility and punishment. This lies in direct contrast to Smilansky who argues that the illusion of moral responsibility is necessary both to make sense of moral actions but also to appease the general public. Pereboom's theory is thus optimistic about renovating our current theories about moral responsibility and punishment in opposition to Smilansky's pessimistic theory which argues that our abilities to make sense of moral responsibility without illusionism are nonexistent.

However, Pereboom's proposal of a modified conversational theory of moral

responsibility also has flaws. I have two objections to Pereboom's modified conversational theory of moral responsibility hard incompatibilism: (1) the conversational theory of moral responsibility rejects the premise that holding an agent morally responsible depends on facts about that agent being morally responsible, and (2) the conversational theory of moral responsibility requires the existence of multiple agents with finely-tuned reactive attitudes.

The first objection turns on my disagreement with the interdependence thesis on which the conversational theory depends. The interdependence thesis, inherent in McKenna's conversational theory of moral responsibility (and in Pereboom's modified theory), says that facts about holding an agent morally responsible are independent of facts about the agent being morally responsible. However, this seems intuitively incorrect to me. There are certain actions that we deem to be immoral regardless of their bearing on the moral community. If this is true, we do regard facts of holding an agent morally responsible to be dependent on facts about being morally responsible.

For example, imagine that Jim likes to torture small animals, and it gives him great pleasure to do so. Let's suppose that torturing small animals has no bearing on the moral community, either because no one finds out or no one is significantly affected by its consequences or both. The interdependence thesis argues that facts about holding Jim morally responsible for torturing small animals do not depend on his being morally responsible for torturing small animals. However, it seems precisely *because* the torturing of small animals is immoral that we regard Jim as morally responsible for his actions: Jim's performing an action with a certain moral charge is a necessary condition of the

moral community holding him responsible for that action. Thus I believe that a weakness in the conversational theory of moral responsibility is that it depends on the interdependence thesis, which seems *prima facie* incorrect.

My second objection to the conversational theory of moral responsibility is that it requires the existence of multiple agents with Strawsonian reactive attitudes. Consider Jim again, but this time he is the only surviving human on Earth. The conversational theory of moral responsibility requires that, for a conversation with moral responsibility exchange to take place, there must exist multiple moral agents to *have* the conversation. We can conceive of a moral community with one moral agent, Jim himself, but for the moral conversation to progress to the second the third stages, there must be a moral agent to perform the address, analyze the response from Jim, and impose the correct moral responsibility ascription.

There are three routes open to the conversational theorist of moral responsibility in response to this objection. The first route is that they can regard non-human animals as moral agents. However, I think most ethicists would be reluctant to do under so under this theory for a few reasons. Under the conversational theory of moral responsibility, it is unclear how a non-human animal could contribute to the moral community at all by performing a moral action. Furthermore, a non-human animal could engage in the second stage, moral address, by expressing its pain or dislike to Jim torturing it, but it is not clear that the animal could engage in the third stage of the conversation either. For these reasons, and because it is not clear that animals have the rationality or capacity to being reason responsive at all, I do not believe that the conversational theorist could argue that

non-human animals are moral agents.

The second route open to the conversational theorist is to argue that Jim, the sole member of the moral community, could have a conversation and moral responsibility exchange *with himself*. However, it is not clear why Jim would engage in moral conversation in the first place, given that his actions have no effect on anyone else other than himself. He has no reason to consider potential victims, his own moral formation, or reconciling nonexistent moral relationships. Furthermore, it is not clear how this conversation could happen in the first place. Even if Jim did have a reason to engage in moral conversation with himself, I do not believe that the moral conversation would achieve any significant end. It is hard to see how or why any moral conversation Jim could have with himself would affect his moral formation or how he currently treats himself as a moral agent.

The third route open to the conversational theorist is to assert that there are no praiseworthy or blameworthy actions without the existence of a moral community. I think this is the most plausible line for the conversational theorist to take. However, I have argued that we are justified in believing in the existence of normative facts due to the non-optional project of moral deliberation. If suddenly all other humans on the planet disappeared but Jim, he would still engage in moral deliberation with regards to various actions: treatment of animals, treatment of the environment, and treatment of himself.⁹⁴ If Jim still engages in moral deliberation and is justified in believing in the existence of normative (and thus moral) facts, certainly he is justified in believing that his actions are

⁹⁴ It is entirely possible that Jim would fail to engage in moral deliberation at all. The main point driving this hypothetical is pure, personal intuition.

still moral or immoral, deserving of the appropriate ascription regardless of the existence of a moral community. Even if these normative facts are socially constructed, they do not lose their moral charge if the moral community disintegrates. Torturing non-human animals is still wrong if there are no other humans around to praise or blame it.

In short I believe that the conversational theory of moral responsibility is unsatisfactory on two counts: (1) it relies on the interdependence thesis, and (2) it requires the existence of a moral community composed of multiple agents with finely-tuned reactive attitudes. For these reasons, I find the conversational theory of moral responsibility lacking, because it is intuitive to me that we regard that facts about holding agents as morally responsible depend on facts about them being morally responsible, and that one can still perform moral and immoral actions that have no bearing on the moral community.

Chapter 5: Conclusion

In this honors project I have performed an analysis of nonaction-centered and action-centered event-causal incompatibilist views. I have argued that event-causal incompatibilist views of free will and determinism are no more philosophically tenable than their compatibilist counterparts. I have shown that these theories offer no increase in proximal control, nor can they answer the problem from moral chanciness. Based on these reasons, I believe that event-causal incompatibilist theories ought to be rejected in favor of more simplistic theories of moral responsibility, namely compatibilist theories.

I have offered a compatibilist theory of free will and determinism of my own, grounding moral responsibility in the existence of normative facts, which are instrumentally indispensable for the intrinsically indispensable project of moral deliberation. The control necessary for free action is captured by guidance control, explicated and defended in the philosophy of compatibilists John Martin Fischer and Mark Ravizza. I have also defended my compatibilist theory against three objections from philosophers Saul Smilansky and Ishtiyaque Haji.

I have offered two alternative, nonstandard views of moral responsibility - illusionism and conversational theory - for those philosophers who do not yet find my compatibilist theory convincing or are unwilling to accept *prima facie* moral realism, on which my compatibilist theory is based. However, these theories also have their flaws, and I have explicated these flaws in the objections herein. Ultimately I believe that these nonstandard views also fail to provide satisfactory accounts of moral responsibility.

If moral responsibility can be salvaged, it must be done by dismantling determinism, or adopting a compatibilist theory. For event-causal incompatibilists that reject agent-causal and nonaction-centered views, the only viable option is adopting a compatibilist theory, for event-causal incompatibilism is closest to compatibilism in terms of scope and application. For the sake of parsimony, the event-causal incompatibilist view ought to abandon their theory for compatibilism.

Bibliography

- Clarke, Randolph. *Libertarian Accounts of Free Will*. Oxford: Oxford Univ. Press: (2014). <http://www.oxfordscholarship.com/oso/public/content/philosophy/019515987X/toc.html>.
- Clarke, Randolph. "Libertarian Views: Critical Survey of Noncausal and Event-Causal Accounts of Free Agency." In Robert H. Kane (ed.), *The Oxford Handbook of Free Will*. Oxford University Press, (2002): 356-385.
- Clarke, Randolph. "Modest libertarianism." *Philosophical Perspectives* 14, no.13 (2000) :21-46.
- Clarke, Randolph and Capes, Justin, "Incompatibilist (Nondeterministic) Theories of Free Will", *The Stanford Encyclopedia of Philosophy* (Spring 2017 Edition), Edward N. Zalta (ed.).
- Enoch, David. *Taking Morality Seriously : A Defense of Robust Realism*. Oxford: Oxford University Press, (2011). Web.
- Fischer, John Martin. "Frankfurt-Type Examples and Semi-Compatibilism." In Robert H. Kane (ed.), *The Oxford Handbook of Free Will*. Oxford University Press, (2002): 281-307.
- Fischer, John Martin. *The Metaphysics of Free Will: A Study of Control*. Oxford: Blackwell.
- Fischer, John Martin and Ravizza, Mark. *Responsibility and Control: A Theory of Moral Responsibility*. Cambridge University Press, (1998).
- Frankfurt, Harry G. "Alternate Possibilities and Moral Responsibility." *The Journal of Philosophy* 66, no. 23 (1969): 829-39. doi:10.2307/2023833.
- Haji, Ishtiyaque. "Alternative Possibilities, Luck, and Moral Responsibility." *The Journal of Ethics* 7, no. 3 (2003): 253-275.
- Mele, Alfred R. *Free Will and Luck*. Oxford: Oxford University Press, (2006). Print.
- Mele, Alfred R., and Oxford University Press. *Autonomous Agents : From Self-Control to Autonomy*. New York: Oxford University Press, (2001). Web.

- Mele, Alfred R. (1999). "Ultimate Responsibility and Dumb Luck." *Social Philosophy and Policy*, 16, no. 2 (1999).
- Pereboom, Derk. *Free Will, Agency, and Meaning in Life*. First edition. Oxford: Oxford University Press, (2014). Web.
- Smart, J. C. C. "Free-Will, Praise and Blame." *Mind, New Series* 70, no. 279 (1961): 291-306.
- Smilansky, Saul. "Free will, Fundamental Dualism, and the Centrality of Illusion." In Robert H. Kane (ed.), *The Oxford Handbook of Free Will*. Oxford University Press, (2002): 489-505.
- Strawson, Galen. "The Impossibility of Moral Responsibility." *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, 75, no. 1/2, Free Will, Determinism, and Moral Responsibility (1994): 5-24.
- Strawson, P F. *Freedom and Resentment, and Other Essays*. London: Methuen [distributed in the USA by Harper & Row, Barnes & Noble Import Division], (1974). Web.
- Timpe, Kevin. "Source Incompatibilism and Its Alternatives." *American Philosophical Quarterly* 44, no. 2 (2007): 143-55.
- Van Inwagen, Peter. "Moral Responsibility, Determinism, and the Ability to do Otherwise." *The Journal of Ethics* 3 (1999): 343. doi:10.1023/A:1009832417573.
- Zimmerman, Michael J. (2002). "Taking Luck Seriously." *Journal of Philosophy* 99, no. 11 (2002): 553-576.