

THE HEART OF LIBERTARIANISM: FUNDAMENTALITY AND THE WILL

Forthcoming in *Social Philosophy and Policy*

It is often claimed that libertarianism offers an unattractive conception of free will and moral responsibility because it renders free agency inexplicable and irrational. This essay aims to show, first, that the soundness of these objections turns on more basic disagreements concerning the ideals of free agency and, second, to develop and motivate a truly libertarian conception of the ideals of free agency. The central contention of the paper is that the heart of libertarians' ideal of free agency is the ideal of *agential fundamentality*.

Key words: free will; moral responsibility; the will; causal theory of action; identification theories; contrastive explanations; rationality

1 Introduction

While libertarianism about free will and moral responsibility has always found supporters, the majority of philosophers see it as a “wretched subterfuge” (to stand Kant’s famous accusation on its head). The main target of attack is libertarians’ insistence that directly free actions for which agents are directly morally responsible must be undetermined. Some argue that undetermined actions cannot be free actions and, consequently, determinism is not only compatible with free will but necessary for it.¹ Others argue that while it may be possible for undetermined actions to be free, the fact that an agent’s action is undetermined does nothing to enhance her control in so acting, and thus indeterminism is superfluous to freedom and responsibility.² Still others have argued that even if libertarianism can solve these conceptual quandaries, it remains indelibly at odds with a sober scientific portrait of humans.³

The present essay considers two further objections that are related but ultimately independent the above worries. These objections share in common the claim that libertarianism renders free agency somehow *undesirable*. That is, these objections aim to show that even if libertarianism can silence the three above-mentioned worries, we still ought to reject the theory because it sets forth an unattractive ideal of free agency. According to *the explanation objection*, libertarianism renders free agency inexplicable: libertarians insist on conditions of free agency

(specifically, the requirement of indeterminism) that are incompatible with the more basic requirement of agential explicability.⁴ According to *the rationality objection*, libertarianism severs or unduly weakens the relationship between the will and reason. By requiring that our free exercises of agency be undetermined, libertarians prevent our agency not only from being determined by the past and laws of nature, but also from being determined by the dictates of reason. But freedom from reason is not freedom “worth wanting.”⁵

I have two aims in this essay. The first is to show that the soundness of these objections turns on more basic disagreements concerning the ideals of free agency: competing visions of the ideal relationship between agency, explanation, and rationality. I argue that the explanation and rationality objections implicitly depend on controversial and to some degree unmotivated ideals of free agency. My second aim is to develop and motivate a truly libertarian conception of the ideals of free agency and show how this conception has the resources to silence the explanation and rationality objections. The central contention of the paper is that the heart of libertarians’ ideal of free agency is the ideal of *agential fundamentality*. What is attractive about the libertarian conception of free agency, I argue, is that this view, and only this view, affords agents a fundamental role in nature.

2 The Will

An interesting feature of the contemporary literature on free will is the conspicuous absence of discussions of the will. The philosophical novice might have expected any discussion of free will to begin with detailed analyses of the will and only then turn to explicate the conditions under which the will is free. However, if philosophers discuss the will at all, it is usually only to set it aside.⁶

Chisholm approvingly cites Locke’s claim that the central issue is not “whether the will be free” but “whether a man be free,”⁷ and Mele notes that he “often cannot tell what authors mean by ‘will’ in

‘free will,’” contending that it is best to define ‘free will’ in terms of ‘free action.’ Mele explains: “one can go about the business of trying to understand free action without worrying about what (the) will is supposed to be. I find that thought liberating.”⁸ While I suspect that there is a grain of truth in Chisholm and Mele’s remarks, it is, I suggest, a mistake to avoid analyzing and making use of the idea of will in our analyses of free will. My aim in this section is to isolate what I take to be the central experience of willing and offer the beginnings of an account of the will. This analysis will serve as the first step toward isolating a truly libertarian ideal of free agency.

The most prominent theory of action in the contemporary literature is the event-causal theory of action.⁹ According to this theory, the causal role of the agent is reducible to the causal role of mental states and events involving the agent. A thief’s robbing a poor box is nothing over and above the thief’s desire for money and his belief that he can best acquire money by robbing the poor box jointly causing his robbing the box. This is a reductionist theory of agency in that it reduces the causal role of the agent to the causal role of his mental states and events.

I want to raise a problem for the event-causal theory of action and explain how this problem may move us toward a nonreductionist theory of agency. My goal is not to provide a decisive refutation of the event-causal theory, but rather to isolate a conception of free agency that I believe is at the heart of libertarianism.¹⁰ Reflection on this problem is, I believe, the best route to appreciating the libertarian ideal.

The problem with the event-causal theory is that it does not appear capable being extended to capture more robust forms of agency, such as self-determination. Cases of motivational conflict reveal the limits of the theory. It is our experience of motivational conflict that gives rise to the idea of will, and it is precisely this experience that the event-causal theory of action cannot account for. I have a familiar kind of motivational conflict in mind. This conflict is an unfortunate and pervasive feature of our agency and has achieved canonization in the conscience of Western philosophical and

religious traditions.¹¹ One of the most powerful presentations of this conflict is set out in Leo Tolstoy's tale of Evgeny Irtenev's sexual obsession with the peasant girl Stepanida.¹² After Evgeny's father passed away, he moved back to his family's estate to set the family affairs in order and try to make their estate profitable again. To preserve his "health," he establishes, with the help of his foreman, a regular "liaison" with Stepanida. Upon his later marrying, he decides to break off all relations with Stepanida and embark on a new course of life. Yet, unexpectedly, Stepanida finds her way back into Evgeny's life when his wife hires some local women to clean their home in preparation for Holy Week. To Evgeny's surprise, he cannot stop thinking about the woman he was sure he had forgotten. Despite his best efforts, Evgeny finds a long-forgotten passion overtaking him. The following passages perceptively recount Evgeny's struggles:

And he turned back to his room, but, before he went five steps, himself not knowing why and on whose order, he glanced around again, so as to see her one more time.¹³

And suddenly passionate lust seared him, clutching his heart like a hand. As if by someone else's alien will, Evgeny glanced around and went after [Stepanida].¹⁴

The main thing was that he felt defeated, that he had no will of his own, and that there was another force that moved him.¹⁵

Evgeny's story is one of a cultivated passion slowly overtaking him. I believe that the event-causal theory of action cannot make sense of this all-too-frequent experience. It is natural enough to suppose that Evgeny's actions that aim at satisfying his lust are caused by his passion. The problem for the event-causal theory is that Evgeny's experience is one in which his motivations are *other* than him. Indeed, it is one in which his motivations work against him. His experience is decidedly not one of self-determination but one of other-determination—namely, determination by his passions.

Frankfurt sought to capture this kind of experience with his notion of alienation. Referring to his well-known case of the unwilling addict (an addict who does not "really want" to be moved

into action by his desire for drugs), Frankfurt writes: “the unwilling addict may meaningfully make the analytically puzzling [statement] that the force moving him to take the drug is a force other than his own.”¹⁶ The mere fact that our bodily movements are caused by mental states is not sufficient for our exercising the power of self-determination, for when we are alienated from our motives that move us into action we fail to exercise this power.

Some worry that Evgeny’s experience is self-deceptive and thus should not be trusted as a guide to the nature of the will.¹⁷ As Shoemaker notes, “the mere fact that we do not like some aspect of ourselves does not make it any less a part of ourselves. We may resign ourselves to it or undertake some course of action designed indirectly to alter or purge it . . . but as long as it remains, it is, for better or worse, part of who we are.”¹⁸ But this misunderstands my and Frankfurt’s claim. I agree that Evgeny’s being alienated from these desires does nothing even to suggest that he is not responsible either for having them or acting on them. Nor does it suggest that these desires are not revelatory of his character. My claim is that it shows his will is not wholly reducible to his motivations, for in this case his motivations work against and without his will.

Frankfurt and others have tried to specify more precisely what it is for an agent to be alienated from a desire or other form of motivation, as well as what it is for an agent to be identified (the opposite of alienation) with a desire or other form of motivation. Most (including Frankfurt himself) have abandoned Frankfurt’s original attempt to analyze identification in terms of second-order volitions, but many other models have been proposed, such as that we are identified with our cares or loves, self-governing policies, evaluative judgments, or the desire to make most sense of ourselves.¹⁹

The problem, I submit, for all these accounts is that our exercises of the power of self-determination are not invariably tied to any of these motivations or any disjunction of these motivations: we exercise the power of self-determination in astonishingly disparate motivational

contexts. We can exercise our power of self-determination without or against our (deepest) cares, without self-governing policies with which we are satisfied, without or against our judgments of what is best, and without or against the desire to make most sense of ourselves. The power of self-determination is a power that is not reducible to the casual role or force of any *other* form of motivation. To take just one example, perhaps the most natural reductionist reply to Evgeny's case is to argue that he failed to be self-determining in his actions not because his causal involvement was exhausted by the causal involvement of his mental states, but because his causal involvement was exhausted by the *wrong* mental states. Evgeny judged that it was best to break off his relations with Stepanida; had these judgments been effective rather than his desires, Evgeny would have been self-determining. This response, however, is shortsighted, for it is possible to perform self-determined perverse actions: an agent can act against his judgment about what is best, not because of weakness but out of defiance. As Watson perceptibly observes, "When it comes right down to it, I might fully 'embrace' a course of action I do not judge best; it may not be thought best, but is fun, or thrilling; one loves doing it, and it's too bad it's not also the best thing to do, but one goes for it without compunction."²⁰

Of course, far more than these cursory comments would be required to show that all identification accounts are inadequate. My aim right now is not prescriptive but diagnostic.²¹ I am trying to ferret out the heart of libertarians' ideals of free agency. Suppose, then, that I am right and the power of self-determination is not reducible to the causal role of our motivations. It seems that the only option left is to envision this power as conferring a power on the agent to be fundamentally causally relevant and not merely derivatively relevant by way of his motivations. Possessing the power of self-determination requires that we possess a *will*: the power to be fundamentally causally involved in deliberation, decision, and overt action. In cases of motivational conflict, we experience our motivations as distinct from ourselves, and in cases where we exercise

strength of will we experience ourselves as distinctive sources of activity—sources that are not wholly reducible to our motivations. These facts, I suggest, constitute the main motivation for the agent-causal theory and the main difficulty for the event-causal theory. How can I be self-determining if the entire range of my agential activities are reducible to mental states that I am not identical to? Taylor nicely expresses this intuition: “If I believe that something not identical to myself was the cause of my behavior—some event wholly external to myself, for instance, or even one internal to myself, such as a nerve impulse, volition, or whatnot—then I cannot regard the behavior as being an act of mine, unless I further believed that I was the cause of that external or internal event.”²² According to the agent-causal theory, we possess the power to be fundamentally or irreducibly causally involved in our agency: what we do is not wholly reducible to what our mental states do.²³ The will, as I conceive it, is the agent-causal power identified by proponents of the agent-causal theory.

Four features of the will need to be highlighted. First, the will is a second-order power.²⁴ It is possible to deliberate, decide, and act without exercising our will. Evgeny’s case provides a particularly powerful example of this possibility, but more mundane cases illustrate the same point. A desire to eat may flit through my mind, quickly followed by the memory that I have decided to have nothing more to eat until dinner, followed by the recognition that I should stick by this judgment, which in turn leads to my refraining from eating. None of these activities requires an exercise of the will, though our freedom with respect to and moral responsibility for them do require that we *could* have exercised our will. In most cases, for example, we have the power to exercise our will to slow down and think matters over more carefully. Steward nicely captures this point:

I am in charge in the way that a government minister is in charge of a department. No minister directly controls all the work of a particular department. . . . But if it becomes

important, a minister can step in to take a more direct interest in a particular matter. She is in control, even though she is not doing all the work.²⁵

An exercise of the will is the exercise of a power that modifies how we exercise other powers. Willing is not an event distinct from choice or action; it is a distinctive way of being involved in choosing and acting. My willing a decision to X consists in my fundamentally causing the decision, and willing an action consists in my fundamentally causing the action. Our will gives us the power to be fundamentally causally involved in the production of our activities. Beings with wills have the power to exert force concerning what they attend to, the intensity and duration of their attention, their decisions and plans, and their execution of these decisions and plans. In exercises of the will the agent himself, and not just his mental states, is causally involved.²⁶

Second, the will is a power to exert force that is, to some degree, independent of our current motivation. The will's power lies in the fact that beings with wills have the power to act in ways that contravene their current strongest motivations, whether these be beliefs, desires, or cares. There is a growing body of psychological literature that supports this claim. According to this work, the will is like a muscle in that it can be fatigued and strengthened.²⁷ These findings suggest that the will is a distinctive factor in how one acts, not wholly reducible to one's other motivations. How one acts depends not only on one's current beliefs, desires, and cares, but also on the strength of one's will.²⁸

Third, following O'Connor, I contend that the agent's motivations structure the probabilities associated with exercises of the will.²⁹ My desire to be as wise as I can when setting out my career goals for the next five years makes it likely that I will exercise my will to sustain deliberation for a lengthier time than usual, weigh competing considerations with serious attention, and solicit advice from mentors. Had this desire been absent, then, all else being equal, I would have been less likely to exercise my will in these ways.

Fourth, following Clarke, I adopt an integrated agent-causal theory, according to which some of the agent's motivations are always joint causes with the agent when he exercises his will.³⁰ Assuming that our thief above exercises his will in deciding to rob the poor box, his decision to rob the poor box is jointly caused by him and his mental states—such as his desires and beliefs—that supported the decision. I understand the relationship between exercises of the will, or the agent-causal power, and the agent's motivations as follows. The agent's motivations behind his decision D are causal contributors to his deciding to D.³¹ The thief desires to steal. This desire increases the likelihood that he will exercise his will to decide to steal. But, we can suppose, he also desires to keep a promise he made to his mother to live a morally good life. This desire increases the likelihood that he will exercise his will to decide to refrain from stealing. As before, suppose he decides to steal. In this case, his desire to steal is a partial cause of his deciding to steal and his desire to keep his promise to his mother is causally inert (at least with respect to this decision).

3 The Heart of Libertarianism

The will is not always free. Freedom of the will requires not only that we have a will, but also the freedom to exercise it in a plurality of ways. Compatibilists and libertarians must part ways at this juncture. Compatibilists contend that the mere fact that one's exercise of the will is necessitated by conditions outside of one's control is no hindrance to freedom. Libertarians demur, contending that free exercises of the will cannot be causally determined.

Agent-causal compatibilists can achieve a degree of agential fundamentality.³² Call this 'synchronic fundamentality.' Suppose Evgeny wills to focus his attention on the deleterious effects of his liaison with Stepanida. This, regardless of determinism, is a fundamental feature of reality: we cannot reduce Evgeny's willing to any other motivation. However, if we imagine that Evgeny's willing was causally determined, there is another sense in which his will is *not* a fundamental feature

of reality: it is not *diachronically fundamental* because it is derivative from the past and laws of nature. Agent-causal libertarians contend that freedom of the will requires synchronic and diachronic agential fundamentality: free exercises of the will are neither reducible to other happenings at the time of action nor wholly derivative from happenings before the time of action.

Free agents are, according to libertarians, then, doubly fundamental. Our free exercises of the will are undetermined, and thus not derivative from our past (including our current motivational composition) and the laws of nature. We find in the resolution of our decisions and actions a fundamental source of new reality—a source that cannot be completely described or explained in terms of other features of reality. It is the irreducibility and indeterminacy of our agency that accords us a fundamental role in nature.

Before returning to the two earlier mentioned objections, I want to draw out some consequences of this conception of free agency.³³ First, free will (so conceived)³⁴ confers the powers of self-governance. A self-governing agent is one who adopts and orders goals, ideals, values, and the means to accomplishing them. Such an agent has not only the power to determine for herself what kind of person to aim at becoming, what kind of activities to invest her time in, and what kind of things to attend to; she also has the power to rank these often competing aims and to determine how she will pursue them and resolve conflicts. We are all partly products of the past, of our heredity, upbringing, and environment. These conditions set limits on how we can pursue life. Free will is not absolute. It is shaped, molded, and limited by numerous factors outside our control. But beings with free will retain control over how they live within these strictures. Their lives are not a mere outflowing of the past and laws of nature, nor are their actions a function of their current motivational dispositions. They can shape and mold their motivational dispositions in ways that run contrary to their current dispositions. It these powers that enable the *self* to govern (within limitations) how his life unfolds.

In addition to the intrinsic value of the powers of self-governance, these powers are also instrumentally valuable because they ground a deep form of moral responsibility. An agent who exercises free will determines for herself what to do. Her choices and actions are not mere functions of her current motivational economy or the past and laws of nature. Her choices are rather a function of her exercises of the power to be fundamentally creative. It would seem, then, that these activities ground desert: the activities and consequences of such exercises belong to the agent in the strictest sense imaginable and thus, depending on the moral valence of the action and the agent's understanding, the agent deserves praise or blame for her activity. While my present aim is not to argue that this libertarian conception of free will is necessary for moral responsibility, it is worth noting that it seems to provide a better candidate for a sufficient condition for moral responsibility than either reductionist or compatibilist models.

4 Reason and Free Will

So goes the libertarian ideal. Compatibilists will reject the significance of diachronic fundamentality and reductionist-minded libertarians and compatibilists will reject the significance of synchronic fundamentality. A full defense of this conception clearly requires further work. But my present concern lies in a different direction. I want to put this conception of free agency to work by showing how it can address the rationality and explanation objections, both of which contend that libertarianism offers a deeply unattractive conception of human agency.

Turning first to the rationality objection, libertarians contend that an agent with free will is synchronically fundamental, and so the work of the will is not in lockstep with the work of reason. Moreover, they contend that an agent with free will is diachronically fundamental, and this means that a free agent is not only free from the past and laws but also free from reason. But, as Hilary Bok claims, "it is not clear why anyone would think [this kind of freedom is] desirable."³⁵ What we

want is not freedom *from* but freedom *within* reason.³⁶ Indeed, even many libertarians see this as a significant problem.³⁷

The worry here concerns the value of free will. Free will entails that we have the power to resist not only our strongest desires, but even the dictates of reason. Is this something of genuine worth? Following Wolf, I understand *reason* to be a power of the mind to form judgments about what is good, desirable, and worthwhile, as well as draw comparisons between competing values in order to determine which is best.³⁸ What, then, is valuable about being able to act contrary to what one judge's best?³⁹

Many who object to the value of free will portray reason as operating wholly before the activities of the will. First one deliberates, weighs the reasons at hand, and then judges which course is best. It is after, and only after, this process that the agent has the freedom to act with or against reason. This view of deliberation is mistaken. The operations of reason are infused with the operations of will. Whether and how long we deliberate, the amount of time spent in gathering reasons before weighing them, the sources of information we go to, the degree of attention we give each consideration, and the amount of effort we exert are often a function of our will. Reason is limited and must be supplemented by will (as well as desire).⁴⁰ Everything we do cannot be preceded by a judgment about what is best. That would simply be too time-consuming. It is therefore to our advantage to be able to act without reason: to act prompted by instinct, habit, and desire without any need to apply our judgment.⁴¹

But the initial worry remains. *When* we exercise our reason and come to judge that some course of action is best, would we not prefer to be the kind of agent for whom acting contrary to our judgment is impossible?⁴² I believe we would, but this question does not isolate the point of disagreement. The relevant question, rather, is: Would it be better to be the kind of agent whose disposition, through no work of his own and for which he deserves no moral praise, is always to

follow the dictates of reason rather than a creature who has free will with respect to whether or not he cultivates this disposition, with the attendant possibility of becoming someone who lives a less than fully rational life? While the answer to the first question seems obvious, the answer to the second question should give us pause. It may well be that a central issue behind the compatibilist/incompatibilist divide concerns a difference of values, incompatibilists contending that it is better to be responsible for coming to be the kind of agent who always acts in accord with reason and compatibilists balking. The rationality objection to the value of free will requires us to answer the second question affirmatively: it is better to be a person who is determined by non-agential factors to have the disposition always to obey reason. But this conception of agency can hardly be *assumed* when criticizing libertarianism, which, at its core, embraces a conflicting ideal of free agency. According to libertarians, the ideal of free agency is agential fundamentality, where agents are synchronically and diachronically fundamental to the flow of their lives. This ideal need not conflict with an ideal of rationality, but it does imply that the ideal of rationality is not absolute. We rightly want to be rational, but not at all costs. Our ideal of rationality, so claim libertarians, should be conditioned by the ideal of our rationality being (to some degree) self-determined, and this very possibility requires that we also have the freedom to reject the dictates of reason. What is desirable about this is not that we can disobey reason *per se*, but that it is fundamentally up to us whether we obey.⁴³

The rationality objection assumes a controversial ideal of free agency. What can be said in favor of this ideal? Why should we prefer to be deprived of the power to disobey reason at the expense of agential fundamentality? Compatibilists might appeal to one (or both) of the following related answers. First, if our acting in accordance with reason is guaranteed by factors outside our control, then we will always act in accordance with reason, and it is a good thing always to act in

accordance with reason. Second, the freedom to act against reason is a liability that opens us up to the undesirable possibility of acting irrationally.⁴⁴

I find both answers wanting. The first answer is insufficient because, while it may well be good always to do what one judges best, that fact does not show that it is better to be determined by factors outside our control always to act in accordance with reason than to have the freedom to act against reason. The second answer seems overblown. The very enterprise of valuing is laden with risk since it is always possible to lose what is of value. Thus, the fact that we can misuse our free will does not show that it would be better to lack it. Moreover, the charge of liability can be exaggerated. What is at stake here is how free the will is. Libertarians contend that freedom of the will concerns even reason: we are free to follow or reject the dictates of reason. To say this is a liability can be misleading, as it suggests that it is a kind of handicap—a force that works against us when we try to align our wills with reason. But this is confused. The supposed “liability” *just is* our freedom of the will. It is up to us whether we obey reason and thus failures to follow reason cannot be viewed as something that happen to us: they are something we freely choose. This is not to deny that we can have *other* liabilities to disobey reason. Addictions and compulsions can be liabilities to disobey reason, but this is because addictions and compulsions are forces working against us in precisely the way our wills cannot be.

These reflections suggest a third reason for the compatibilist ideal: the libertarian ideal is too risky. Even granting that the freedom to act irrationally is not a liability along the lines of an addiction or compulsion, it still opens us to failure. Compatibilists could concede that it is preferable that we come to have the disposition always to act rationally as a result of our own agency and yet that, due to risk, it is all-things-considered better not to be subject to this risk.⁴⁵

The attractiveness of each ideal seems partly a function of the perspective we take up. Suppose Ann has become a perfectly rational creature and looking back on her life wonders: Would

it have been better to have had the power to reject reason and yet, despite this, to have made myself into an ideally rationally being? Retrospectively, it is quite easy to feel the force of the libertarian ideal and embrace an affirmative answer. Consider instead Jeff who is at the beginning of his life, not knowing whether he will attain the ideal of rationality and wonders: Would it be best to have the power to reject rationality with its attendant risks? Prospectively, it is quite easy to feel the force of the compatibilist ideal and embrace a negative answer.⁴⁶

It is hard to know exactly how to settle this dispute, but my aim here is less ambitious. My aim is not show that the libertarian ideal is preferable but that it is attractive and offers the resources for dismantling the rationality objection. To see this let us consider a particularly powerful version of the rationality objection developed by Wolf.⁴⁷ To convince us that the freedom to act against reason does nothing to confer greater responsibility on an agent, she asks us to consider an agent, Ann, who comes upon a child drowning in a lake. Ann is a strong swimmer out for a leisurely walk without any pressing obligations. As far as she can tell, she is the only one to witness the child's peril. She is the kind of person who, without any conscious deliberation, recognizes that there is decisive reason to save the child and given this recognition it was not possible for her to have made any other choice or performed any other action than to save the child. She makes this decision and saves the child from drowning.⁴⁸

Now consider a second agent, Jeff, who is identical to Ann in every relevant respect except that his choices are not causally determined by his recognition of decisive reasons. Jeff also recognizes that there is decisive reason to save the child and decides to save the child. However, Jeff, unlike Ann, retained the power to refrain from so deciding. He possessed the power to act against his own reason.

Does Jeff bear more responsibility than Ann for saving the child? Wolf thinks not: "Each agent acts, in fact, in exactly the way we want agents to act and for exactly the motives we want

agents to have. At first glance it seems that each agent deserves as much credit as an agent in this situation can possibly deserve.”⁴⁹ This is too quick though. First, the mere fact that agents perform the same kind of action on the basis of the same kind of motives hardly establishes sameness of responsibility. This is clearest in the case of an adolescent and adult acting wrongly. Even if they perform the same action for the same motives, we (all else being equal) hold the adult more responsible than the adolescent.

Second, issues of history play a significant role. It is quite natural upon hearing the story to think not only that Jeff deserves no more praise than Ann, but that Ann actually deserves *more* praise than Jeff. The naturalness of this response stems from our tendency to assume that adults bear responsibility for the kind of persons they have become. Ann is the kind of person for whom the child’s distress compels her to save him. She is so committed to the good that she can do nothing else but pursue it. This is indeed valuable. But Wolf’s argument requires us to stipulate that neither Ann nor Jeff bear responsibility for their characters, for without this stipulation libertarians are not committed to Jeff’s being more praiseworthy for saving the child than Ann. This stipulation, however, weakens her case since it affords libertarians a plausible explanation of why Jeff is more responsible, an explanation that appeals to their ideal of free agency: Jeff is more praiseworthy because his choice to save the child is synchronically and diachronically fundamental. It is a free choice, expressive of his powers of self-governance. He is more responsible than Ann because her activities are derivative; they are a function of the past and laws. Ann is fortuitous in becoming the kind of person who would act in this noble way, but while her actions are admirable, she does not deserve praise for them. As Reid noted, saying “*He was good because he could not do otherwise...* if understood literally and strictly, is not the praise of [the agent], but of his constitution, which was no more the work of [the agent], than his existence.”⁵⁰

The rationality objection turns on a controversial ideal of free agency in which we purchase perfect rationality at the cost of agential fundamentality. When this conception is replaced with libertarians' ideal, the force of the objection evaporates. I suspect that much of the objection's force stems from a misformulation of the issue, which is not whether it is desirable to be the kind of agent who always acts according to reason—for on this question libertarians and compatibilists can agree—but whether always obeying reason is worth the sacrifice of agential fundamentality. The ideal of agential fundamentality implies that the ideal of rationality is not absolute: agential fundamentality requires the freedom to act irrationally.

5 Fundamentality and Explanation

The second worry about the desirability of libertarianism concerns agential explicability. At times it has been suggested that undetermined action is essentially irrational action, or, at the very least, cannot be given a reasons-explanation.⁵¹ But this, as most now recognize, is false. The central features of a successful reasons-explanation are that the explanation makes intelligible to us why the agent acted as he did and that the features appealed to which render the action intelligible play a causal role in bringing about the action. Agent-causal libertarianism can satisfy both desiderata. Consider again our thief, Jones, who, let us suppose, successfully resists the temptation to rob the poor box at a local parish. We might ask: Why did Jones decide to refrain from stealing? To this libertarians can answer: he suddenly recalled that he had promised his mother on her death bed to live a good life and he recognized that stealing from the poor is inconsistent with this promise. Moreover, these reasons played a causal role in bringing about his decision to refrain. Jones's reasons meet both the above conditions: they make his decision intelligible and they play a causal role in bringing about his decision. So libertarians can, in a straightforward way, secure reasons-explanations for undetermined actions.

But there is another kind of explanation that has played a more prominent role in objections to libertarianism, and it is less clear that this kind of explanation is available for undetermined actions. This is *contrastive explanation*—explanation of why one thing rather than another occurred. For any event *E*, there are countless other contrasts to it that might lead us to ask why it occurred rather than some other event. For example, we might ask why Jones decided to refrain from stealing rather than join a monastery, or why he made this decision rather than deciding to run a marathon. The contrast we decide to focus on is determined by our interests. However, libertarianism entails that for every directly free action *X* an agent performs at time *t*, it was possible—given the past and laws up until *t*—that he *Y* instead (where ‘*Y*’ may simply be refraining from doing anything). It is precisely this contrast that critics of libertarianism have focused on. Let us call this *the core contrast*: the agent’s *X*-ing rather than *Y*-ing.⁵² In Jones’s case, the core contrast is his deciding to refrain rather than deciding to steal. According to the explanation objection, libertarianism cannot secure a key desideratum of free agency: namely, that our free actions and choices be contrastively explicable.

Some critics of libertarianism use this conclusion as a premise in an argument that aims to show that since undetermined actions cannot be contrastively explained, undetermined actions are lucky in a way that is incompatible with their being free actions.⁵³ However, others have suggested the lack of agential inexplicability itself is devastating to libertarians.⁵⁴ The idea is that embedded in our ordinary practices of moral responsibility is asking for and giving contrastive explanations of our actions. Libertarianism, so the objection goes, provides an unattractive ideal of free agency because it undermines the accuracy of these folk practices, implying that free actions cannot be contrastively explained. It is the second contention that I aim to respond to here.

The explanation objection, I contend, turns on an equivocation.⁵⁵ If the meaning of ‘contrastive explanation’ is given a type-2 reading, then we can give contrastive explanations of free

choices. If the meaning of ‘contrastive explanation’ is given a type-1 reading, then while it is true that no free choices can be contrastively explained, that is a virtue of the account, not a vice. I use ‘type-1/type-2’ rather than ‘complete/incomplete’ or ‘full/partial’ to avoid the idea that there is some sense in which type-2 contrastive explanations fail to completely or fully explain.

The sense of a type-2 contrastive explanation is inspired by Hitchcock’s influential model of contrastive explanation.⁵⁶ According to Hitchcock’s model, we can contrastively explain why Jones decided to refrain rather than steal simply by citing his reasons that favored refraining. Jones decided to refrain rather than steal because he remembered that he promised his mother to live a good life. This reason causally contributed to his decision to refrain, and it would not have causally contributed to his decision to steal had he made *that* decision instead. It is this difference that allows us to cite the reason as an explanation of why the thief decided to refrain rather than steal.

Following Hitchcock—albeit with a slight modification—if *E* causally contributes to the occurrence of *C*, and would not have causally contributed to the occurrence of not-*C* had not-*C* occurred, then citing *E* explains why *C* rather than not-*C*—even though *C*’s occurrence was undetermined and even if *C*’s occurrence is overall less probable than *C*’s nonoccurrence.⁵⁷

Turning to a different case, suppose that Jane is deciding between vacationing in Hawaii and Colorado. She has strong reasons supporting each decision, never judges that one option is better than the other, and finally decides to vacation in Hawaii.⁵⁸ Clarke contends that we cannot give even a poor contrastive explanation of Jane’s choice.⁵⁹ Rejecting Hitchcock’s model, he endorses Lipton’s Difference Condition: “To explain why *P* rather than *Q*, we must cite a causal difference between *P* and not-*Q*, consisting of a cause of *P* and the absence of a corresponding event in the history of not-*Q*.”⁶⁰ A corresponding event is “something that would bear the same relation to *Q* as the cause of *P* bears to *P*.”⁶¹ What exactly this condition comes to is unclear, but it is clear that Lipton’s account yields different verdicts about the availability of contrastive explanations than

Hitchcock's account. In Jane's case, she has reasons that support her decision to go to Hawaii and reasons that support her decision to go to Colorado. Moreover, the reasons are too close for her to judge that one option is better. Lipton's account entails that we cannot contrastively explain Jane's decision because there is no causal difference between the actual causal process culminating in her decision to go to Hawaii and the possible causal process culminating in her decision to go to Colorado. In either case, the decision is caused by the reasons that bear *the same kind of relation* (i.e., the causal relation or favoring relation) to the choice they favor.

But this seems like the wrong verdict. Suppose you are Jane's friend, you know that she was deliberating between vacationing in Hawaii and Colorado, and that she has decided to go to Hawaii. Suppose also that up to this point you know nothing about the details of her deliberation. You ask: "Why did you decide to go to Hawaii rather than Colorado?" Jane responds: "Well, I was really torn about where to go, but I really wanted to go snorkeling and surfing, and so decided to go to Hawaii." So long as these reasons causally contributed to her decision to vacation in Hawaii and would not have causally contributed to her deciding to vacation in Colorado, then it seems to me that we have a genuine contrastive explanation and a *good* one at that. We have it without making any further assumptions about whether Jane judged that it was best to go to Hawaii, whether there is any other causal difference between her decision in the actual world and her decision in some other possible world, or whether the causal process leading to her choice was deterministic or nondeterministic.

Contrastive reasons-explanations aim to make intelligible why an agent chose one option rather than another. We can provide such explanations by citing reasons that causally contributed to our making one choice and would not have causally contributed to our making the alternative choice. Suppose Jane responded to your question by saying, "I really wanted to get away." This fails to provide even a poor contrastive explanation of her choice not because it fails to cite a causal

difference in Lipton's sense, but rather because the cited condition causally contributes to her choosing to go to Hawaii and *would have* causally contributed to her choosing to go to Colorado had she made that choice instead. In seeking a contrastive explanation of Jane's choice, we are seeking information that makes intelligible why she chose Hawaii rather than Colorado. The availability of such explanation merely requires that there be reasons that favor Hawaii rather than Colorado and that these reasons causally contributed to deciding to go to Hawaii and would not have causally contributed to deciding to vacation in Colorado. In citing such reasons, we make Jane's choice to go to Hawaii rather than Colorado contrastively intelligible.

Suppose that receiving this answer from Jane, you respond: "But you also wanted to hike in the Rockies." Maybe you really wanted her to go to Colorado and are frustrated with her choice and so you press her: "So, why did you choose Hawaii rather than Colorado?" To this Jane might naturally respond by pointing out other distinctively attractive features of Hawaii. But, of course, you might respond in kind by pointing out further distinctive features of Colorado that were also attractive to her. After pressing her long enough, she will say what we all say when pressed like this: "I just decided that was what I wanted to do." Some might be tempted to draw the conclusion from this that, in the end, we cannot *really* contrastively explain her choice after all. But I see no reason to draw that conclusion. Jane has already given an adequate contrastive explanation of her choice. What is happening is that you are probing to see if there is *further* explanatorily relevant information. And of course there usually is. For almost all of our decisions, we never cite all of the reasons (and nonreasons) that contributed to our making the choice. The list would simply be too long. Because of this, it makes perfect sense to probe deeper. But if we press long enough, at some point we will run out of explanatorily relevant information. While Jane's original response provides an adequate contrastive explanation, it would be inappropriate for her simply to offer it again in response to your subsequent questions. What you wanted is further information. But the

fact that there is no more information to give does not imply that there was never any information to give.

Hence, if ‘contrastive explanation’ means type-2 contrastive explanation, then the explanation objection fails: we can give contrastive explanations of free choices. There is, however, another sense of ‘contrastive explanation’ that the objector may have in mind. There cannot be a contrastive explanation of an agent’s directly free action that cites a factor prior to and independent of the agent’s activity that guarantees that he acts one way rather than another. In this type-1 sense, directly free actions, according to libertarianism, are inexplicable. The question now is whether the necessary absence of this kind of explanation is a vice of libertarianism.

Part of our common sense framework for thinking about free will and moral responsibility is that free actions are actions done for reasons. Any account of freedom that severs all ties to reason is untenable. Moreover, it is a frequent part of our explanatory practice to ask and give answers to questions such as: “Why did he choose this rather than that?” We ask such questions about both mundane facts of life and life-changing choices. Therefore, it is plausible to think that if a theory cannot accommodate the possibility of contrastive explanations at all, that is at least a mark against it. But libertarianism can account for the kinds of explanations we give in our ordinary, interpersonal practices. In such cases, we cite the reasons that causally contributed to our choice and that would not have causally contributed to the alternative choice. This seems enough to ground our ordinary practices. However, proponents of the explanation objection contend that “real” or “true” contrastive explanations of an action must cite factors that *guarantee* that the action occur rather than not. If we are to give a contrastive reasons-explanation of Jane’s decision to go to Hawaii, then—so the thought goes—it must be possible to cite reasons that guarantee that, assuming the presence of other uncontested necessary conditions (e.g., Jane remains alive), she will make this decision rather than not. Libertarianism cannot accommodate such explanations, but this,

I maintain, is a virtue of the theory. That is, there are certain kinds of contrastive explanations that we should not want to be available for all our actions. If all our actions are guaranteed by our prior motivations and other factors, then we do not fundamentally determine what we do. The issue here is not whether what we try or intend to do will guarantee what we will do. A case can be made for the desirability of that connection. Rather, proponents of the explanation objection insist that it is desirable that our agency be completely determined by *nonagential* factors. If type-1 contrastive explanations are available for all our actions, then everything we do is causally determined by what we do not do: our agency is wholly determined by nonagency. Are proponents of the explanation objection right to think that this kind of explanation desirable?

It is important to remember that the question before us is not whether exercises of free agency are compatible with such explanations but whether the availability of such explanations is desirable: should we want our agency to be wholly (even if not exclusively) determined by nonagential factors? Libertarians think not. If all our actions are guaranteed by our prior motivations and other factors, then we do not fundamentally determine what we do. Agential fundamentality, the libertarian ideal of free agency, requires diachronic fundamentality, which requires the impossibility of type-1 contrastive explanations of an agents' directly free choices.

The explanation objection presupposes the falsity of the libertarian ideal of agential fundamentality, and thus it is unsurprising that libertarians consider the objection misguided. The very thing missing that the proponent of the explanation objection finds objectionable is the very thing whose absence libertarians find attractive. But, more to the point, without a defense of the ideal of free agency motivating the type-1 explanation objection, the objection is unmotivated. So what is the ideal behind this objection? What is attractive about the possibility of giving type-1 contrastive explanations of an agent's free exercise of the will? As far as I can tell, defenders of the explanation objection point to the desirability of agential explicability: we should be able to explain

the intentional and free activity of agents—indeed, we should be able to contrastively explain their activities. But libertarians can, as we have seen, provide contrastive explanations of free actions. The general ideal of agential explicability does not provide support for the explanation objection. Rather, the proponent of the explanation objection needs to explain what is attractive about the more specific possibility of giving type-1 contrastive explanations of free actions. As far as I know, no proponent of this objection has addressed this question.

I suspect the reason for this lacuna is that there is simply no plausible way to address it. The attraction of libertarianism is that it promises to accord us, free agents, a fundamental role in the production of free actions. If libertarianism is true, then our directly free choices are not the inevitable consequence of the past and laws of nature. Part of our fundamentality as free agents requires that our free exercises of the will not be explicable in certain senses in terms of *other* features of reality. If our exercises of free will were made inevitable by the past and laws of nature, if there were conditions prior to and independent of our choices that guaranteed that we make the choices we make, then we would not be fundamental but merely derivative sources of activity. The very absence of this kind of explanation—the very absence of prior motivations that guarantee that the agent will make one decision rather than another—is a requirement for agential fundamentality. The impossibility of type-1 contrastive explanations on libertarianism is an attraction of the view, not an embarrassment.⁶²

Conclusion

My aim has been to isolate and motivate the ideal of free agency behind libertarianism and thereby reveal that the rationality and explanation objections assume the falsity or unattractiveness of this ideal. Perhaps the assumption behind these objection is correct, but it is an assumption that cannot legitimately be made when offering objections against libertarianism. However, even if I have

succeeded in this aim, this does not show that libertarianism is true. Much would be required to show that. But I do hope to have shown that libertarianism is motivated by a distinctive ideal of free agency that may well be worth wanting.⁶³

¹ See, e.g., R. E. Hobart, "Free Will as Involving Determination and Inconceivable without It," *Mind* 43 (1934): 1–27; A. J. Ayer, "Freedom and Necessity," in *Philosophical Essays* (New York: St. Martin's Press, 1954), 3–20.

² See, e.g., Gary Watson, "Free Action and Free Will," *Mind* 96 (1987): 154–72.

³ See, e.g., Manuel Vargas, "Libertarianism and Skepticism about Free Will: Some Arguments against Both," *Philosophical Topics* 32 (2004): 403–26; Derk Pereboom, *Free Will, Agency, and Meaning in Life* (Oxford: Oxford University Press, 2014), Ch. 3.

⁴ See, e.g., Alfred Mele, *Free Will and Luck* (New York: Oxford University Press, 2006), Ch. 3; Ishtiyaque Haji, "Reason, Responsibility, and Free Will: Reply to My Critics," *Journal of Ethics* 16 (2012): 175–209.

⁵ Daniel Dennett, "On Giving Libertarians What They Say They Want," in *Brain Storms: Philosophical Essays on Mind and Psychology* (Montgomery: Bradford Books, 1978), 286–99; Susan Wolf, *Freedom within Reason* (New York: Oxford University Press, 1990); Richard Double, *The Non-reality of Free Will* (New York: Oxford University Press, 1991); Hilary Bok, *Freedom and Responsibility* (Princeton, NJ: Princeton University Press, 1998).

⁶ Important exceptions include Harry Frankfurt, "Freedom of the Will and the Concept of a Person," *Journal of Philosophy* 68 (1971): 5–20; Robert Kane, *The Significance of Free Will* (New York: Oxford University Press, 1996); Gary Watson, "The Work of the Will" in *Agency and Answerability: Selected Essays* (New York: Oxford University Press, 2004), 123–59.

⁷ Roderick M. Chisholm, "Human Freedom and the Self," in Gary Watson, ed. *Free Will* (New York: Oxford University Press, 2003), 34.

⁸ Alfred Mele, *Free Will and Luck*, 17.

⁹ For a lucid defense of this theory see John Bishop, *Natural Agency: An Essay on the Causal Theory of Action* (New York: Cambridge University Press, 1989).

¹⁰ Not only will compatibilists reject this ideal of agency, but so also will libertarians who embrace (something like) the event-causal theory of action (e.g., Kane, *Significance of Free Will*, and Laura Waddell Ekstrom, *Free Will: A Philosophical Study* (Boulder, CO: Westview Press, 2000)).

¹¹ In Plato's example of the Charioteer (*Phaedrus*) and St. Paul's description of the effects of sin (Romans 7), for example.

¹² Leo Tolstoy, "The Devil," in *The Death of Ivan Ilyich and Other Stories*, trans. Richard Pevear and Larissa Volokhonsky (New York: Random House Inc., 1888 [2009]), 164–208.

¹³ Tolstoy, "The Devil," 183.

¹⁴ *Ibid.*, 188.

¹⁵ *Ibid.*, 189.

¹⁶ Frankfurt, "Freedom of the Will and the Concept of a Person," 18.

¹⁷ Michael McKenna and David Shoemaker both raised this worry in conversation.

¹⁸ David Shoemaker, "Caring, Identification, and Agency," *Ethics* 114 (2003): 115

¹⁹ See, respectively, Harry Frankfurt, "Autonomy, Necessity, and Love," in H. F. Fulda and R. P. Horstmann, ed., *Vernunftbegriffe in der Moderne: Stuttgarter Hegel-Kongress 1993* (Stuttgart: Klett-Cotta, 1994); David Shoemaker, "Caring, Identification, and Agency"; Chandra Sripada, "Self-Expression: A Deep Self Theory of Moral Responsibility," *Philosophical Studies* 173 (2016): 1203–32; Michael Bratman, "Reflection, Planning, and Temporally Extended Agency," *Philosophical Review* 109 (2000): 35–61; Gary Watson, "Free Agency," *Journal of Philosophy* 72 (1975): 205–20; J. David Velleman, "What Happens When Someone Acts?" *Mind* 101 (1992): 461–81.

²⁰ Watson, "Free Action and Free Will," 168.

²¹ I have, however, offered extensive discussion and criticism elsewhere. See Christopher Evan Franklin, "Self Determination, Self-Transformation, and the Case of Jean Valjean: A Problem for Velleman," *Philosophical Studies* 172 (2015): 2591–98; "If Anyone Should Be an Agent-Causalist, then Everyone Should Be an Agent-Causalist," *Mind* 125 (2016): 1101–31; "Bratman on Identity over Time and Identification at a Time," *Philosophical Explorations* 20 (2017): 1–14; "Cares, Agency, and Identification Reductionism," forthcoming in *Pacific Philosophical Quarterly*; and *A Minimal Libertarianism: Free Will and the Promise of Reduction* (New York: Oxford University Press, 2018), Ch. 7.

²² Richard Taylor, *Metaphysics* (Englewood Cliffs: Prentice-Hall, 1974) 55; cf. Franklin, "If Anyone Should Be an Agent-Causalist, then Everyone Should be an Agent-Causalist," 1119–27 and Franklin, *A Minimal Libertarianism*, Ch. 7.

²³ For key explanations and defenses of the agent-causal theory see Timothy O'Connor, *Persons and Causes: The Metaphysics of Free Will* (New York: Oxford University Press, 2000); Randolph Clarke, *Libertarian Accounts of Free Will* (New York: Oxford University Press, 2003), Chs. 8–10; Meghan Griffith "Why Agent-Caused Actions Are Not

Lucky,” *American Philosophical Quarterly* 47 (2000): 43–56; Helen Steward, *A Metaphysics for Freedom* (Oxford: Oxford University Press, 2012).

²⁴ The will’s being a second-order power explains both how we are continuous and discontinuous with other animals. Our agency shares much in common with other mammals since much of what we do is wholly reducible to the causal activities of our mental states, just like animals. What distinguishes us is the development of a power to modify how we exercise these powers in a way that is not merely a function of our current motivational states.

²⁵ Steward, *A Metaphysics for Freedom*, 51

²⁶ An important implication of this account of free will is that we can perform actions over which we have free will even though we do not exercise free will in so acting. That is, we can perform actions that we do not play a fundamental causal role in bringing about even though we could have. While such actions will not be instances of the exercise of free will according to my account, we can still say we had free will over such an action and thus the action is free. Thanks to Mark Balaguer for raising this issue.

²⁷ See, e.g., Mark Muraven and Roy F. Baumeister, “Self-regulation and Depletion of Limited Resources: Does Self-control Resemble a Muscle?” *Psychological Bulletin* (2000) 126: 247–59.

²⁸ Richard Holton, *Willing, Wanting, Waiting* (New York: Oxford University Press, 2009), Ch. 6. This of course does mean that your will operates wholly independent of your motivations either. It is plausible to think that our motivations set some bounds on what options we can will. Much more needs to be said here, but the point I am making about the will is not that it operates wholly independent of our motivations, but that it is not wholly dependent on our motivations. Thanks to Mark Balaguer for raising these issues.

²⁹ Timothy O’Connor, “Agent-Causal Power,” in Toby Handfield, ed., *Dispositions and Causes* (New York: Oxford University Press, 2009), 189–214.

³⁰ Clarke, *Libertarian Accounts of Free Will*, Ch. 8.

³¹ *Ibid.*, 135–36

³² See Ned Markosian, “A Compatibilist Version of the Theory of Agent Causation,” *Pacific Philosophical Quarterly* 80 (1999): 257–77; Ned Markosian, “Agent Causation as the Solution to all the Compatibilist’s Problems,” *Philosophical Studies* 157 (2012): 383–98; Dana K. Nelkin, *Making Sense of Freedom and Responsibility* (New York: Oxford University Press, 2011).

³³ The next two paragraphs draw on Franklin, *A Minimal Libertarianism*, Ch. 4.

³⁴ From hereon I will drop the qualification ‘so conceived.’

³⁵ Bok, *Freedom and Responsibility*, 48. This is one of the varieties of free will that Dennett claims is not worth wanting: “The libertarian could not have wanted to place the indeterminism *at the end* of the agent’s assessment and deliberation” (Dennett, “On Giving Libertarians What They Say They Want,” 51). Dennett thinks such a desire would be “insane.”

³⁶ Wolf, *Freedom within Reason*, Ch. 3.

³⁷ Ekstrom, *Free Will*, 104–106; Steward, *A Metaphysics for Freedom*, 127–33.

³⁸ Wolf, *Freedom within Reason*, 53–54. For simplicity I will assume that there is always a uniquely best option.

³⁹ As Mark Balaguer and Justin Capes have pointed out to me in conversation, one might dismiss this objection outright: After all, what does the desirability of an account have to do with the truth of the account? I don’t think the rationality and explanation objections can be that easily dismissed. Frankfurt (“Freedom of the Will and the Concept of a Person,”) contends, rightly to my mind, that an adequate theory of free will must explain why we find free will attractive, something worth wanting. Thus, while in general the fact that we find an account unattractive has little bearing on the truth of the account, it is quite relevant if we are trying to give an account of something we find attractive. Thanks to Dana Nelkin for some helpful ideas on how to think about these issues.

⁴⁰ Cf. Thomas Reid, *Essays on the Active Powers of the Human Mind* (Cambridge: MIT Press, 1788 [1969]), III.i–ii.

⁴¹ Holton, *Willing, Wanting, Waiting*, Chs. 4–6.

⁴² Cf. Mele, *Free Will and Luck*, 111–17.

⁴³ An additional way to respond, as Michael McKenna and Bob Kane have pointed to me, is to call into question the prevalence of all-things-considered judgments. It is very often not the case that when we make a crucial life-setting decision we can tell which option is best. While I am convinced that this is correct, I will not take up this response. Rather, I want to show that even if we assume that such all-things-considered judgments are quite prevalent, there is reason to be attracted to the libertarian ideal.

⁴⁴ Watson, “Work of the Will,” 129–30.

⁴⁵ I owe this point to Justin Coates

⁴⁶ I owe this point to Matt Talbert.

⁴⁷ Wolf, *Freedom within Reason*, Ch. 3.

⁴⁸ *Ibid.*, 58–59.

⁴⁹ *Ibid.*, 59.

⁵⁰ Reid, *Essays on the Active Powers of the Human Mind*, IV.i.

⁵¹ A. J. Ayer, "Freedom and Necessity."

⁵² For some directly free actions, it will be possible that the agent have done one of a number of different things, given the past and laws of nature, in which case there will be multiple core contrasts. I ignore this complication since it does not affect the substance of my argument.

⁵³ See, e.g., Mele, *Free Will and Luck*, Ch. 3; Haji, "Reason, Responsibility, and Free Will: Reply to My Critics."

⁵⁴ C. D. Broad, "Determinism, Indeterminism, and Libertarianism," in *Ethics and the History of Philosophy* (London: Routledge and Kegan Paul, 1952), 195–217; Rebekah L. H. Rice, "Agent-Causation and Acting for Reasons," *American Philosophical Quarterly* 48 (2011): 333–46.

⁵⁵ Aspects of the following argument are borrowed from Franklin, *A Minimal Libertarianism*, Ch. 5.

⁵⁶ Christopher Hitchcock, "Contrastive Explanation and the Demon of Determinism," *British Journal of the Philosophy of Science* 50 (1999): 585–612; "Contrastive Explanations," in *Contrastivism in Philosophy: New Perspectives*, ed. Martijn Blaauw (London: Routledge 2012), 11–34.

⁵⁷ Hitchcock, "Contrastive Explanation and the Demon of Determinism," 599. The slight modification concerns the fact that Hitchcock states his account in terms of probability raising rather than causal contribution. His account runs as follows: if *E* raises the probability of *C* more than not-*C*, then citing *E* explains why-*C* rather than not *C*—even though *C*'s occurrence was undetermined and even if *C*'s occurrence is overall less probable than *C*'s nonoccurrence (Hitchcock, "Contrastive Explanation and the Demon of Determinism," 602).

⁵⁸ This case is taken from Kane, *Significance of Free Will*, 28.

⁵⁹ Clarke, *Libertarian Accounts of Free Will*, 45–46.

⁶⁰ Peter Lipton, *Inference to the Best Explanation* (London: Routledge, 1991), 43.

⁶¹ Lipton *Inference to the Best Explanation*, 43.

⁶² Fred Miller has pointed out to me in conversation that one might think that my response to the rationality and explanation objection are individually successful but jointly incompatible. In responding to the explanation objection, I argue that instances of free will can be given contrastive reasons-explanations of type-1. In response to the rationality objection, I argue that free will requires the possibility of acting irrationally. The key to seeing the compatibility between these claims is that type-1 contrastive reasons-explanation do not require that one act for the best reasons, but

just for some reasons. Thus, we can give a reasons-explanation of an irrational action so long as the agent had at least some reasons for acting against his best reasons.

⁶³ Thanks to Mark Balaguer, Justin Coates, Bob Kane, Ben Mitchell-Yellin, Garrett Pendergraft, Micah Quigley, Philip Swenson, Matt Talbert, and Neal Tognazzini for helpful comments. Earlier drafts of this paper were presented at the 2013 Tennessee Value and Agency Conference, University of Tennessee, and Moral Responsibility: The Next Generation Workshop, University of Arizona. I am grateful to the conference participants, especially E. J. Coffman and Manuel Vargas, for their helpful comments. I owe a special thanks to Michael McKenna for his careful comments on earlier drafts of this paper and hard and excellent work to make this workshop and issue of this journal possible.