

CONSTITUTIVE MORAL LUCK AND STRAWSON'S ARGUMENT FOR THE
IMPOSSIBILITY OF MORAL RESPONSIBILITY

Robert J. Hartman
University of Gothenburg
roberthartman122@gmail.com

(Forthcoming in the *Journal of the American Philosophical Association*)

Abstract: Galen Strawson's Basic Argument is that because self-creation is required to be truly morally responsible and self-creation is impossible, it is impossible to be truly morally responsible for anything. I contend that the Basic Argument is unpersuasive and unsound. First, I argue that the moral luck debate shows that the self-creation requirement appears to be contradicted and supported by various parts of our commonsense ideas about true moral responsibility, and that this ambivalence undermines the only reason that Strawson gives for the self-creation requirement. Second, I argue that the self-creation requirement is so demanding that either it is an implausible requirement for a species of true moral responsibility that we take ourselves to have or it is a plausible requirement of a species of true moral responsibility that we have never taken ourselves to have. Third, I explain that Strawson overgeneralizes from instances of constitutive luck that obviously undermine true moral responsibility to all kinds of constitutive luck.

Galen Strawson (1994; 2002; 2010; 2011) formulates in various ways his *Basic Argument* for the claim that true or ultimate moral responsibility is impossible, whether or not causal determinism obtains. Here is his briefest formulation: Because self-creation is required for being truly morally responsible and self-creation is impossible, it is impossible to be truly morally responsible for anything. Let us consider my simplified reconstruction of Strawson's most informative formulation:

Reasons Premise: An agent S's intentionally performing an action A for which she might be truly morally responsible is explained by certain features of her mental constitution MC—namely, certain reasons for acting.¹

Responsibility Premise: S is truly morally responsible for an intentional action A only if S is truly morally responsible for the parts of her MC that explain her performing A, and S is truly morally responsible for her MC only if S is truly morally responsible for an earlier action A₁ in which S intentionally and successfully brought about those parts of her MC.

¹ Strawson (2010: 28, cf. 2010: 34) notes that even if some free actions are not performed for reasons, the "vast majority" of them "must" be performed for reasons.

Iteration Premise: S is truly morally responsible for A by way of MC and A_1 as previously described only if S has performed an infinite number of even earlier free actions. (After all, S is truly morally responsible for A_1 only if S is truly morally responsible for the parts of her MC_1 that explain her performing A_1 , and S is truly morally responsible for her MC_1 only if S is truly morally responsible for an even earlier action A_2 in which S intentionally and successfully brought about those parts of her MC_1 . S is truly morally responsible for A_2 only if S is truly morally responsible for the parts of her MC_2 that explain her performing A_2 , and S is truly morally responsible for her MC_2 only if S is truly morally responsible for an even earlier action A_3 in which S intentionally and successfully brought about those parts of her MC_2 . S is truly morally responsible for A_3 only if ...)

Impossibility Premise: It is impossible for finite beings like us to have performed an infinite number of past actions.

Conclusion: It is impossible for finite beings like us to be truly morally responsible for anything.

What these formulations have in common is that being truly morally responsible for something requires a kind of self-determination that is impossible for us to have.

Strawson (1994: 5, 8) has lamented that his Basic Argument is underappreciated. Recently, however, the Basic Argument, or something relevantly like it, has received broad endorsement (Hendrickson 2007; Istvan 2011; Kershner 2015; Kment 2017; Nagel 1986, p. 118; Smilansky 2000, Ch. 4; Waller 2011: 19-42). Others have noted their sympathy for it (Henderson 2014: 328; Mickelson 2015; 2017).

Those of us, however, who believe that being truly morally responsible is at least possible must find fault with one of these premises. The Reasons Premise is uncontentious. After all, there is widespread agreement that actions for which we are truly morally responsible are actions that are performed for reasons. The Impossibility Premise is obviously true, and the Iteration Premise is just an entailment of the Responsibility Premise. So, the possibilist should take aim at the Responsibility Premise.

Possibilists have provided at least three reasons to doubt the Responsibility Premise. One reason is that agent-causal libertarianism is possible, and that agents with agent-causal power have

the right kind of control over their actions to be truly morally responsible for them without also being truly morally responsible for features of their mental constitutions that explain their actions (Clarke 1997; Corabi 2017; O'Connor 2011: 320-321; Pereboom 2001: 65-68; Tucker 2007).

Another reason involves a G. E. Moore Shift against the Responsibility Premise: It is more plausible that possibly we are truly morally responsible agents than that the Responsibility Premise is true (Coffman 2015: 113-115; Fischer 2006: 112). Yet another reason is that that the Responsibility Premise is implausible, because it is demanding in a way outstrips commonsense requirements of true moral responsibility (Clarke 2005; Fischer 2006; Kane 1996; McKenna 2008; Mele 1995: 221-230; Russell 2008; Wolf 1990; 2015).

My argument is situated roughly in the third camp. In what follows, I offer three arguments for the claim that the Responsibility Premise is unjustified and implausible, none of which appeal to agent-causal libertarianism or a G. E. Moore Shift.

In the first section, I briefly explicate Strawson's (1994: 6-7) idea of true or ultimate moral responsibility. In the second section, I argue that the moral luck debate shows that the Responsibility Premise is contradicted by part of our commonsense ideas about moral responsibility and appears to be supported by another part, and that this ambivalence undercuts the only reason that Strawson gives for thinking that the Responsibility Premise is true. In the third section, I argue that the Responsibility Premise has no foundation in common sense. The Responsibility Premise is so demanding that either it is an implausible requirement for the true moral responsibility that we take ourselves to have or it is a plausible requirement of a species of true moral responsibility that we have never taken ourselves to have. In the fourth section, I suggest that the intuitive attractiveness of the Responsibility Premise is the result of a generalization from instances of constitutive luck that obviously undermine true moral responsibility to all

instances of constitutive luck. But this is an overgeneralization precisely because certain kinds of constitutive luck are *benign* with respect to true morally responsible agency and other kinds at least partially *enable* it. The joint force of these arguments is that the Responsibility Premise is unjustified and implausible and that the Basic Argument is unpersuasive and unsound.

True Moral Responsibility

Strawson is specific about the kind of true moral responsibility that the Basic Argument targets. He (1994: 6-7; cf. 2002: 451-453; 2010: 2) writes,

What sort of ‘true’ moral responsibility is being said to be both impossible and widely believed in?

An old story is very helpful in clarifying this question. This is the story of heaven and hell. As I understand it, true moral responsibility is responsibility of such a kind that, if we have it, then it *makes sense*, at least, to suppose that it could be just to punish some of us with (eternal) torment in hell and reward others with (eternal) bliss in heaven. The stress on the words ‘makes sense’ is important, for one certainly does not have to believe in any version of the story of heaven and hell in order to understand the notion of true moral responsibility that it is being used to illustrate. ... The story of heaven and hell is useful simply because it illustrates, in a peculiarly vivid way, the *kind* of absolute or ultimate accountability or responsibility that many have supposed themselves to have, and that many do still suppose themselves to have. It very clearly expresses its scope and force (*italics in original*).

In other words, the “old story” need not be true for Strawson’s purposes. It merely illustrates the scope and force of the widely believed in conception responsibility that his argument targets.

Strawson (2002: 452; cf. 2010: 2) appeals to heaven-or-hell responsibility to illustrate that the desert at issue is basic in the following way: “URD [ultimate responsibility-or-deservingness] is responsibility and desert of such a kind that it can exist if and only if punishment and reward can be fair or just without having any pragmatic justification, or indeed any justification that appeals to the notion of distributive justice.” Strawson (2010: 2) calls the above quotation a “less dramatic but equally effective way of characterizing true moral responsibility” as the heaven-or-hell story. Thus, heaven and hell are used merely to illustrate that what a person deserves is based on only backward-

looking considerations.² But then, reference to heaven and hell is not essential to characterizing the nature of true moral responsibility (Corabi 2017; Istvan 2011: 401). And since several philosophers have rightly noted that appeal to heaven and hell gratuitously introduces confusion into the scope and force of the kind of responsibility at issue (Clarke 2005: 21-22; Fischer and Tognazzini 2011: 25-30; Levy 2011: 3), I consider true moral responsibility as a kind of basic desert and refer to it without reference to heaven or hell.

One more clarification about true moral responsibility is worth making at this point. Strawson (2010: 15) contends that although compatibilist conceptions of moral responsibility such as reasons-responsive views can satisfy “many of the demands of our ordinary notion of freedom,” they cannot satisfy the demand for the kind of “self-determination” that he believes is required for true moral responsibility.³ As Strawson (2010: 15) writes, “[T]he function of the qualifiers ‘true’ and ‘truly’ is simply to keep this point clear: whatever it may be, compatibilist responsibility is not true responsibility in the present sense.” I neither rehearse nor assess Strawson’s (2010: 107-147; cf. 1994: 16-17) arguments for these claims. For dialectical purposes, I grant that true moral responsibility is *incompatibilist* to show that my arguments have traction against the Basic Argument even if Strawson’s arguments against compatibilism succeed. Nevertheless, my arguments do not presuppose incompatibilism; compatibilists can use them too.

In the next section, I undercut Strawson’s justification for the Responsibility Premise by situating it within the broader debate about moral luck.

² Forward-looking considerations can impact whether it is morally permissible to give someone what she deserves. For a nice discussion on the ethics of blaming the blameworthy, see Todd (forthcoming).

³ Philosophers have recently distinguished between different kinds of moral responsibility including attributability, answerability, and accountability (Shoemaker 2015). The kind of moral responsibility that Strawson has in mind is accountability.

The Ambivalence Argument from Moral Luck

Strawson (1994: 12; cf. 2002: 454, 445) asserts that the Responsibility Premise is a part of common sense: “[W]e find, semi-dormant in common thought, an implicit idea that true or ultimate moral responsibility for what one does somehow involves responsibility for how one is.” Strawson (1994: 11; cf. 1994: 17-18; 2002: 454) gives “some examples of people who have accepted that some sort of true or ultimate responsibility for the way one is is a necessary condition of true or ultimate moral responsibility for the way one acts.” His list includes Robert Kane, Immanuel Kant, and Jean-Paul Sartre among others. Thus, Strawson thinks that the Responsibility Premise is a commonsense requirement for true moral responsibility, and he illustrates that claim by appealing to the philosophy of Kane, Kant, and Sartre.⁴

But it is not obvious whether common moral intuitions in fact justify the Responsibility Premise, and we can see why by putting the Basic Argument into the broader dialectic of the problem of moral luck, which is something that philosophers have neglected to do.⁵

The problem of moral luck is that there is a contradiction in our commonsense ideas about true moral responsibility (Hartman 2017: 1). In one strand of our thinking, we believe that a person can become more praiseworthy or blameworthy by luck or in a way influenced by luck. To make this idea concrete, consider some cases of luck in results, circumstance, and constitution:⁶

⁴ Strawson’s claim about Kane and Kant is contentious. Kane (1996: 74-75; 2007: 22) denies that all free actions require tracing, which implies the denial of the first conjunct in the Responsibility Premise (see also Clarke 2005: 16n3). Additionally, it is not obvious whether Kant endorses the Responsibility Premise; for discussion of this issue, see Hartman (manuscript-c).

⁵ There are a few exceptions to this claim including E.J. Coffman and Paul Russell. My use of the moral luck debate, however, differs from Coffman’s (2015: 104-115) use, because he does not appeal to it *as* a response to the Basic Argument. Additionally, my use differs significantly from Russell’s (2008: 319-324) use, because he notes that *compatibilists* should not agree with Strawson that background luck is problematic (see also Hartman 2017: 42-59; Russell 2017); my arguments in this paper do not rely on compatibilism.

⁶ Typically, the problem of moral luck is restricted to conflicting intuitions about cases of resultant, circumstantial, and constitutive luck, because this restriction makes this puzzle obviously distinct from more familiar puzzles about causal determinism, or causal luck (Hartman 2017: 4-9).

Resultant Luck Case: Two agents drive recklessly around a curb in the same way, and one but not the other kills a pedestrian (Nagel 1979: 29).

Circumstantial Luck Case: Two judges would freely take a bribe if one were offered. By luck of the courthouse draw, only one judge is offered a bribe, and so only one judge takes a bribe (Thomson 1989: 214).

Constitutive Luck Case: Two citizens would freely help a beggar if they had a good upbringing, but they were habituated differently. The citizen with good habituation stops to help the beggar, and the citizen with bad habituation ignores the beggar.

In each case, luck is the salient difference between the two agents. After all, the spatial location of the pedestrian is outside of each driver's control; being offered a bribe is outside of each judge's control; and being a recipient of good habituation is outside of each citizen's control. In our everyday lives, however, we make moral judgments that imply that the killer driver is more blameworthy than the merely reckless driver; the bribe taker is more blameworthy than the mere would-be bribe taker; and the helper is more praiseworthy than the mere would-be helper (Nagel 1979).⁷ This is because we believe that certain kinds of luck in results, circumstance, and constitution can positively affect praiseworthiness and blameworthiness. Call the intuitions that represent this part of common sense *the moral luck intuition*.

Nevertheless, the moral luck intuition contradicts another feature of our ordinary thinking about true moral responsibility captured in this *Control Principle*: A person's praiseworthiness and blameworthiness is restricted to what is within her control, and her praiseworthiness and blameworthiness cannot be affected by that which is not within her control—whether it be the results of her actions, the circumstances she faces, or features of her constitution (Zimmerman 2002: 559; see also Nagel 1979). Bernard Williams (1985: 194) explains it this way: “There is pressure within it [our ordinary conception of morality] to require a voluntariness that will be total

⁷ By the phrase ‘S is more blameworthy than S*,’ I mean that S deserves *more blame* than S*, but perhaps it is also true that S is *more deserving* of blame than S*.

and will cut through character and psychological or social determination, and allocate blame and responsibility on the ultimately fair basis of the agent's own contribution, no more and no less." True attributions of praiseworthiness and blameworthiness reflect only a person's pure or luck-free agency. So then, the Control Principle implies that the drivers are equally blameworthy, the judges are equally blameworthy, and the citizens are equally praiseworthy, because the salient difference between the agents in each case pair is luck and luck is irrelevant to true moral responsibility. Call the intuitions that represent this part of common sense *the luck-free intuition*.

Our commonsense ideas about true moral responsibility, then, appear to include both the moral luck and luck-free intuitions. A prominent way to understand the problem of moral luck is that it is not clear which intuition is correct. Thomas Nagel (1979: 34), who is one of the founders of the contemporary moral luck debate, leaves the problem of moral luck as a "paradox," because both parts of common sense appear to him to be undeniable.⁸

Let us relate the problem of moral luck to the Basic Argument, which Strawson (2000: 151; cf. 2002: 449) himself frames in terms of luck at a certain point: "One can put the point (somewhat contentiously) by saying that in the final analysis the way you are is, in every last detail, a matter of luck—good or bad."

If the moral luck intuition is right, then it rules out the truth of the Responsibility Premise, because it implies that an agent's constitutive luck—that is, features of the agent's constitution that she did not bring about—may explain why she performs an action for which she *is* truly morally responsible. For example, the citizen is praiseworthy for helping the beggar even though his lucky constitution explains why he chooses to help. But the Responsibility Premise implies that the citizen is not praiseworthy for helping the beggar, because, on that view, she is not truly morally

⁸ The other one is Williams (1981).

responsible for even part of the habituation that explains why she helps the beggar. Thus, at least part of our commonsense ideas about true moral responsibility contradict the Responsibility Premise.

If the luck-free intuition tracks the truth, then it would support the idea that the Responsibility Premise is true. The Control Principle implies that what is not within the agent's control, such as non-voluntarily acquired parts of her constitutive mental life, cannot affect her true moral responsibility for the action, and the Responsibility Premise implies that the features of a person's constitutive mental life that saliently explain the action for which she is truly morally responsible *must* be constitutive properties that she has intentionally and successfully brought about by previous actions. Thus, at least some of our commonsense ideas about true moral responsibility support the Responsibility Premise, and this fact helps explain why the Responsibility Premise is intuitively attractive.

The upshot is that our moral intuitions do not obviously justify the Responsibility Premise, because part of common sense implies that the Responsibility Premise is false and another part of common sense appears to support it.⁹ For this reason, it is not enough for Strawson merely to appeal to our commonsense ideas about true moral responsibility to justify the Responsibility Premise.¹⁰ He needs to tell us why *only* the luck-free intuition is right. As his presentation of the

⁹ In various places, Strawson also claims that our intuitions are muddled. We believe that the Responsibility Premise is true, but we also experience feelings of being morally responsible due to the phenomenology of our deliberating and choosing (1994: 9; cf. 2002: 455-459). We intuit that "true self-determination," which is the idea embodied in the Responsibility Premise, "can look very odd. But it is also an extremely natural claim. It is at least as natural as it is odd" (2010: 46).

¹⁰ Susan Hurley (2000: 248) makes this same point outside of the moral luck debate: "But given that intuitions conflict about whether responsibility is regressive, why doesn't the impossibility of regressive responsibility count in favor of a nonregressive view of responsibility rather than in favor of eliminating responsibility? This question has no *obvious* answer . . . Neither position should simply be taken for granted" (italics in original). My presentation of the idea is an improvement precisely because the moral luck debate illustrates the nature and scope of our conflicting intuitions. Consideration of the moral luck debate can also help us to see the way in which the Responsibility Premise lacks commonsense credentials; I argue for this in the next section.

Basic Argument currently stands, however, Strawson provides us with no all-things-considered reason to believe that the Responsibility Premise is true. The Ambivalence Argument from Moral Luck, then, shows that the Basic Argument is not persuasive, because the most crucial premise in the Basic Argument suffers from a stalemate of intuitions.¹¹

One might object that the moral luck intuition cannot track *true* moral responsibility, because the moral luck intuition is compatibilist and true moral responsibility is incompatibilist. But the objector would be mistaken. The claim that certain kinds of luck in results, circumstance, and constitution *can* affect true moral responsibility is consistent with both compatibilism *and* incompatibilism about true moral responsibility and causal determinism. The moral luck intuition is consistent with incompatibilism, because there is an important difference between cases of luck in circumstance and constitution, on the one hand, and luck in causal determination—that is, *causal luck*—on the other.¹² The difference is that causal luck is sufficient to bring about the action that it affects, but circumstantial and constitutive luck are mere causal contributors and thus are not sufficient to bring about the action that they affect. Thus, even if certain kinds of circumstantial and constitutive luck can affect true moral responsibility, it does not follow necessarily that the same is the case for casual luck. I should note, however, there is an interesting relationship between the moral luck intuition and compatibilism. In my view, the moral luck intuition provides defeasible evidence for compatibilism, because if luck in results, circumstance, and constitution can positively affect true moral responsibility, then there is at least some reason to think that causal luck can also positively affect true moral responsibility.

¹¹ In Hartman (2017), I argue that the moral luck intuition is correct. If my arguments there are sound, then the intuitions that imply the falsehood of the Responsibility Premise are vindicated. I do not, however, rely on those arguments in this paper.

¹² I assume the lack of control account of luck in this paper; see Hartman (2017: 23-31) for a defense. Some philosophers have additional conditions on luck that preclude casual determinism from being a kind of luck (see Levy 2011: 40).

The Too Demanding Argument

I argue now that even the Ambivalence Argument from Moral Luck is too generous, because the Responsibility Premise is too demanding to capture adequately even part of our commonsense ideas about true moral responsibility.

Someone who is sympathetic to the Basic Argument is likely to object that this line of reasoning inevitably begs the question, because it is *of course* true that the Responsibility Premise is too demanding. This is why it is *impossible* to be truly morally responsible for anything.

Nevertheless, my Too Demanding Argument does not beg this question. It does not involve the following G. E. Moore Shift: It is more plausible that possibly we are truly morally responsible agents than that the Responsibility Premise is true. My Too Demanding Argument also does not amount to this argument: Because it is plausible that we are truly morally responsible for something and the Responsibility Premise implies we cannot be truly morally responsible for anything, the Responsibility Premise is implausible. Rather, the idea behind the Too Demanding Argument is that once we explore sufficiently how demanding the *requirements* of the Responsibility Premise are, we will see that our commonsense ideas about true moral responsibility do not support those requirements. Subsequently, I explain why it is an understandable mistake to believe that the Responsibility Premise is based in common sense in part by scrutinizing the Control Principle, which was introduced in the last section.

Consider, first, that the Responsibility Premise implies that an agent must have chosen the part of her mental constitution that explains her action to be truly morally responsible for the action. For Strawson (1994: 6; 2010: 29-45), the salient part of an agent's mental constitution is the agent's reasons for acting, because, as the Reasons Premise makes clear, the focus of the Basic Argument is intentional actions for which agents can give reasons. The kind of reason for action

that Strawson (2010: 24) has in mind is the internalist belief/desire pair kind. So, we can put the Responsibility Premise in this way: An agent S is truly morally responsible for her action only if she is truly morally responsible for at least part of her beliefs and desires that crucially explain it, and S is truly morally responsible for at least part of those beliefs and desires only if S is truly morally responsible for an earlier action in which S *intentionally* and *successfully* brought about her having those beliefs and desires. As Strawson (2010: 24) writes,

But to be truly responsible for how one is, mentally speaking, in certain respects, one must have chosen to be the way one is, mentally speaking, in certain respects. It is not merely that one must have caused oneself to be the way one is, mentally speaking; that is not sufficient for true moral responsibility. One must have consciously and explicitly chosen to be the way one is, mentally speaking, in certain respects, at least, and one must have succeeded in bringing it about that one is that way.

True moral responsibility, then, requires that a person has consciously, explicitly, and successfully chosen in the past one's present reasons for action.

There are, however, familiar and weighty reasons to think that we cannot consciously, explicitly, and successfully choose our beliefs and desires relevant to free actions (cf. Strawson 2010: 37; Peels 2017: 52-88).

Suppose, however, that we could choose our beliefs and desires in the strong way that Strawson requires to ground to true moral responsibility. Could we, then, be truly morally responsible agents? The answer is clearly 'no' because there is no way to stop the infinite regress of such choices since human beings live only about 130 years (cf. Strawson 2010: 42).

Furthermore, even if we were *infinite beings* who have performed an infinite number of self-creating past actions, it still does not follow that we could be truly morally responsible agents, because the conditions in the Responsibility and Iteration Premises are mere *necessary* conditions on true moral responsibility. As it happens, an infinite being has no advantage for satisfying Strawson's conditions of true moral responsibility. As Strawson (2010: 50) writes, "Not even God could be

truly self-determining as to his motives for action; nor therefore could he be truly responsible for what he did in any ultimate way.” Why should we think that? True moral responsibility requires a buck-stopping action that introduces it in the first place. Strawson (2011) puts it this way: “[T]o be truly or ultimately responsible ... there has to be, and cannot be, a starting point in the series of acts of bringing it about that one has a certain nature—a starting point that constitutes an act of ultimate self-origination.” Thus, having performed an infinite number of actions is no help for being truly morally responsible for something.¹³

Perhaps the best description that Strawson gives of a regress-stopping action that introduces true moral responsibility to go up the causal chain is the *causa sui* (Strawson 1994: 5). The *causa sui* is the ultimate cause of who she is and what she does. Nevertheless, a *causa sui* is not merely metaphysically impossible but is *unintelligible*. As Strawson (2010: 50) himself writes: “One cannot ultimately choose one’s psychological condition; we cannot describe this. It’s not simply that we don’t know how to, but that we know we cannot.” The *causa sui* must both (i) bring about the *foundational* action to *become* truly morally responsible for who she is and (ii) *be* truly morally responsible for who she is *prior* to the foundational action. In other words, the *causa sui* must act to become truly morally responsible for who she is at a time when she is not already truly morally responsible for who she is, and, for her to perform that buck-stopping action, she must already have been truly morally responsible for who she is. Thus, it is *logically* impossible, and not merely *metaphysically* impossible, to be truly morally responsible for something.

Strawson’s requirements on true moral responsibility are, thus, extremely demanding. These descriptions should help us to see that the conditions of the Responsibility Premise are not recognizable in common sense. An openminded agnostic about the soundness of the Basic

¹³ Why, then, formulate the Basic Argument in a way that makes it appear that having performed an infinite number of free actions might help one to be truly morally responsible? The Basic Argument is designed merely to show that true moral responsibility is impossible. It is not designed to tell us how it might be possible.

Argument would, I think, agree. She would inchoately think that we need to be able to create ourselves *in some sense* to be truly morally responsible for our actions. But she would reasonably think that some weaker sense of self-creation is all that is required for true morally responsible agency, because the weaker sense is the only kind of self-creation to be found in common sense.¹⁴

Why, then, do so philosophers believe that the Responsibility Premise, or something relevantly like it, is part of our commonsense conception of morality? One explanation is that they are erroneously extrapolating from ideas that are found in common sense.

We find in common sense the following

Weak Control Principle: An agent is truly morally responsible only for what is within her control.

And one might think that the Weak Control Principle implies the

Strong Control Principle: What an agent is truly morally responsible for cannot be affected by that which is outside of her control.

In other words, the Strong Control Principle is the idea that true moral responsibility requires *total control*: “An agent has total control over X only if for any factor *f* which is a causal contributor to X and which is such that if *f* were not to occur, then X would not occur, the agent has control over *f*” (Fischer 2006: 116). As I argued in the previous section, the Strong Control Principle implies the Responsibility Premise. So then, if the Weak Control Principle implies the Strong Control Principle and if the Strong Control Principle implies the Responsibility Premise, it would follow that the Weak Control Principle implies the Responsibility Premise. The Responsibility Premise would,

¹⁴ Michael McKenna (2008: 202-203) suggests that common sense contains multiple concepts of moral responsibility—and one of them is Strawson’s conception. When we recognize that one of our concepts has an incoherent requirement for being morally responsible, McKenna asserts that we should *not* let it guide our philosophical theorizing; we should rather let the coherent strands of our thinking “inform our philosophical theorizing.” I am arguing that Strawson’s conception is *not* part of commonsense thinking.

then, have a foundation in common sense, because the Weak Control Principle is a commonsense principle.

But the Weak Control Principle does not imply the Strong Control Principle. It is, after all, widely acknowledged that the Strong Control Principle is too demanding to properly describe a requirement for true moral responsibility (cf. Coffman 2015: 103; Fischer 2006: 115-116; Hartman 2017: 49). For an agent to satisfy the Strong Control Principle and so to have total control over an action, she would, for example, need to have had control over a meteorite's failing to demolish the earth or the sun's not flickering out, because these background conditions must have occurred for the agent to perform the action (Fischer 2006: 116). In fact, even philosophers who embrace the *luck-free intuition* agree that the Strong Control Principle is too strong (Levy 2011: 5; Zimmerman 1987: 377-378). One upshot is that we have a clear example of a too demanding style argument that does not beg any questions. Another upshot is that the Responsibility Premise cannot get its commonsense credentials from the Strong Control Principle precisely because the Strong Control Principle lacks such credentials.

Nevertheless, it is important to see that the Responsibility Premise need not depend on the Strong Control Principle for its merits, because the Responsibility Premise is weaker than the Strong Control Principle. As Michael Anthony Istvan Jr. (2011: 405-406) rightly contends, the Responsibility Premise does not require total control over an action to be truly morally responsible for it; rather, the Responsibility Premise requires only that the agent be truly morally responsible for the *mere part* of her mental constitution that explains her action (cf. Strawson 2002: 445).

One might, however, think that the Responsibility Premise receives its commonsense credentials directly from the Weak Control Principle. But this does not seem to be the case either, because the Weak Control Principle obviously implies *only* one conjunct of the Responsibility

Premise. The Weak Control Principle, the idea that an agent is truly morally responsible only for what is within her control, plausibly implies that a person is truly morally responsible for features of her mental constitution only if she had performed an action that brought about that mental constitution (or perhaps if she had that constitution and chose to omit performing actions that would change it). Nevertheless, at the very least *it is not clear* that the Weak Control Principle implies the contentious part of the Responsibility Premise—namely, that a person is truly morally responsible for an action only if she is truly morally responsible for the parts of her mental states that explain her performing that action. Whether or not the Weak Control Principle implies that part of the Responsibility Premise depends on whether one can have true moral responsibility-level control over her action without having such responsibility-level control over the features of her mental constitution that explain why she performs that action.

I contend that the Weak Control Principle does *not* imply the contentious part of the Responsibility Premise. To defend this claim, I briefly sketch a commonsense picture of the relation between true moral responsibility for character and action. At the time when a person *first* has the relevant capacities for true morally responsible agency such as moderate reasons-responsiveness, normative competence, and/or the ability to do otherwise, she performs actions that causally contribute to who she is even though she is not truly morally responsible for those actions. Nevertheless, after she has performed enough of these character-shaping actions for which she is not truly morally responsible, she would be causally responsible for who she is to a sufficient degree such that she can be truly morally responsible for her *actions* to some *marginal degree* (cf. Cyr forthcoming). In this way, true moral responsibility is not *transferred* from earlier actions like on Strawson's picture, but rather it *emerges* from the right kind of non-responsibility conditions. The more of these actions that she performs for which she is truly morally responsible to some degree, the more she becomes truly morally responsible for who she is to some degree. And as she

becomes more truly morally responsible for who she is, she becomes more truly morally responsible for what she does until—again, under felicitous conditions—she crosses a threshold and becomes a full morally responsible agent. The important point is that the commonsense idea of true moral responsibility involves a kind of moral emergence, and this point accommodates Strawson’s idea that self-creation is relevant in some way to true moral responsibility for actions. If this emergence picture belongs to common sense as I claim, it implies that Strawson’s transfer approach deviates from commonsense ideas about true morally responsible agency, and so the Responsibility Premise lacks a foundation in common sense.

We should conclude, then, that the Responsibility Premise is so demanding that either it is false or otiose. If we construe the Too Demanding Argument to show that the Responsibility Premise is false, then the Responsibility Premise is an *implausible* requirement for a species of true moral responsibility that we take ourselves to have. This horn of the dilemma amounts to a refutation of the Basic Argument. If we construe the Too Demanding Argument to show that the Responsibility Premise is otiose, the Responsibility Premise is a plausible requirement for a species of the true moral responsibility that *we have never taken ourselves to have*. This horn of the dilemma shows that the Basic Argument has missed its target, and so is irrelevant. The target of the Basic Argument is a conception of true moral responsibility that is “widely believed in” (Strawson 1994: 6) and “most people ordinarily and unreflectively suppose themselves to possess” (Strawson 2010: 26). Either way, the Basic Argument does not yield a pessimistic conclusion. In the next section, I offer my main error theory of the Responsibility Premise.

Error Theory

If the Responsibility Premise lacks support from common sense, how else might we explain its appeal? I think that its intuitive attractiveness comes from an overgeneralization. I first explain the generalization, and then explain why it is a bad one.

When philosophers think about luck in constitutive properties, they immediately reflect on bad cases in which an agent's constitutive luck obviously undermines her true moral responsibility. Some standard examples include severe emotional trauma, bad formative circumstances, systematic conditioning, and mental illness. One uniting feature of these examples is that the agent does not voluntarily have that kind of history or condition, and these histories and conditions preclude the kind of control required to be truly morally responsible for an action that issues from them.

Consider a famous example from Susan Wolf (1987: 63-64):

JoJo is the favorite son of Jo the First, an evil and sadistic dictator of a small, undeveloped country. Because of his father's special feelings for the boy, JoJo is given a special education and is allowed to accompany his father and observe his daily routine. In light of this treatment, it is not surprising that little JoJo takes his father as a role model and develops values very much like Dad's. As an adult, he does many of the same sorts of things his father did, including sending people to prison or to death or to torture chambers on the basis of a whim. ... In light of JoJo's heritage and upbringing—both of which he was powerless to control—it is dubious at best that he should be regarded as responsible for what he does.

Wolf is correct that we tend to judge that JoJo is not truly morally responsible for his actions (or at least that JoJo is *not fully* truly morally responsible), because JoJo does not have the right kind of control over himself to be truly morally responsible for his actions due to systematic conditioning.¹⁵

¹⁵ David Faraci and David Shoemaker (2010; 2014) use experimental philosophy to argue that most people lack Wolf's intuition that JoJo is not blameworthy to any degree. They instead intuit that JoJo's blameworthiness is only mitigated. For a response, see Michelle Ciburri (2014). For my purposes, it is not important which of these philosophers is correct. All parties agree that JoJo is *not fully* blameworthy for his actions, and this mitigation is all that I need for the error theory. In fact, as Andrew Khoury pointed out to me, I do not even need the claim that JoJo is *not fully*

The generalization, then, proceeds *from* the claim that certain kinds of constitutive mental properties outside of an agent's control (that result from severe emotional trauma, bad formative circumstances, systematic conditioning, and mental illness) undermine true moral responsibility *to* the claim that all kinds of constitutive mental properties outside of an agent's control undermine true moral responsibility. That is, because standard excusing and exempting cases of constitutive luck undermine true moral responsibility, so do other cases of constitutive luck.

Strawson, of course, does not explicitly generalize in this way. In fact, Strawson offers no real argument for the Responsibility Premise but rather asserts that it is to be found in our commonsense ideas about true moral responsibility (Clarke 2005: 19-20; Fischer 2006: 118). Nevertheless, there is reason to think that this generalization may be in the background of his thinking—or at least in the thinking of others who intuit that the Responsibility Premise is plausible—because Nagel's (1979) skeptical argument from luck is a more general version of the Basic Argument and Nagel makes a generalization of this kind.¹⁶

Nagel's skeptical argument aims to eliminate from true moral responsibility not only the luck that frames a person's mental constitution but also the luck that comes before it and after it. A rough summary of Nagel's (1979) skeptical argument is as follows:

Prior to reflection it is intuitively plausible that people cannot be morally assessed for what is not their fault, or for what is due to factors beyond their control (p. 25).

Everything seems to result from the combined influence of factors, antecedent and posterior to action, that are not within the agent's control (p. 35).

[Thus,] the area of genuine agency, and therefore of legitimate moral judgment, seems to shrink under this scrutiny to an extensionless point (p. 35).¹⁷

blameworthy for his actions—but only the claim that anyone who is inclined to find the Basic Argument attractive would believe that JoJo is not fully blameworthy for his action.

¹⁶ As Talbert (2015: 75) points out in passing, “[Galen] Strawson’s argument is ... closely related to skeptical worries raised by moral luck, particularly by constitutive moral luck.”

¹⁷ Recall that Nagel does not endorse this argument but leaves it as a paradox.

What reason does Nagel give for thinking that factors outside of an agent's control cannot affect that for which she is truly morally responsible? Nagel (1979) offers this generalization: "the broad range of external influences [luck in results, circumstance, constitution, and causal determinism] seems on close examination to undermine moral assessment as surely as does the narrower range of familiar excusing conditions" (p. 26) including "involuntary movement, physical force, and ignorance of circumstances" (p. 25). So, Nagel begins with the explanation that various excuses outside of the agent's control undermine true moral responsibility, and then generalizes that all instances of luck—that is, factors beyond the agent's control—undermine true moral responsibility.

Let us return to the *Strawsonian generalization*, which is merely a convenient label for the generalization that I have outlined. I contend that this generalization is a bad one, because there are two broad kinds of constitutive luck that are relevantly different from the previously described excuses and exemptions.

We may call the first kind *responsibility-enabling constitutive luck*; this kind of constitutive luck at least partially enables true morally responsible agency (cf. Hartman 2017: 94-95).¹⁸ It furnishes its agent with the broad range of reason-giving cognitive abilities and reason-responsive volitional abilities required to have what R. Jay Wallace (1994: 157-159) calls "reflective self-control."¹⁹ The cognitive part includes at least the ability to "grasp and apply moral reasons," and the volitional part includes at least the "power to control or regulate his behavior in light of such reasons" (Wallace 1994: 157). All agents who enjoy reflective self-control are necessarily recipients of this kind of constitutive luck, because they cannot make *free* choices to acquire this kind of constitutive make-up precisely because such powers are required to act freely. Many compatibilists *and* libertarians

¹⁸ One might think that this name is question-begging. It is not. As I mention at the end of this section, it is epistemically possible that the constitutive luck that is a necessary condition for true morally responsible agency is also sufficient to undermine true morally responsible agency.

¹⁹ Wolf calls this basic condition "sanity" (1987: 381) and "normative competence" (1990: 124).

recognize that reflective self-control is a necessary condition for morally responsible agency (cf. Fischer and Ravizza 1998; Franklin 2018; Hartman 2017; Wallace 1994; Wolf 1987; 1990).

Responsibility-enabling constitutive luck differs from the kind of constitutive luck that excuses and exempts, because those excuses and exemptions diminish or damage at least some of the capacities required for reflective self-control. JoJo, for example, comes to acquire and endorse his father's values through systematic conditioning of a kind that cuts him off from adequately understanding the weight that moral reasons should have in governing his behavior. As a result, he is not morally responsible (or not fully so) for sending people off to the torture chambers on a whim (cf. Mason 2015: 3055). Another familiar example is Robert Harris who was abused badly by both of his parents all throughout childhood (Watson 1993: 130-139), and this kind of severe emotional trauma (and perhaps psychopathy) almost entirely smothered his ability to see the weight that the moral reasons should have in his deliberations, which is why he is plausibly not morally accountable (or not fully so) for the horrendous double murder he perpetrated (cf. Mason 2015: 3053; Wolf 2011: 334). For convenience, refer to these standard excusing and exempting cases of constitutive luck as *responsibility-undermining constitutive luck*.²⁰

In addition to responsibility-enabling constitutive luck, there is a second kind of constitutive luck that differs from responsibility-undermining constitutive luck, and we may call it *responsibility-banal constitutive luck*. As the name suggests, responsibility-banal constitutive luck is a garden variety part of the agent's mental constitution that is not required for reflective self-control but also does not undermine it. This class of constitutive properties outside of one's control includes a broad range of mental dispositions that contribute to explaining why an agent does what

²⁰ Shoemaker (2015: 191-203) offers an alternative explanation of our reactions to these cases, and a crucial part of his explanation is that it is not the agent's constitutive *luck* that mitigates blameworthiness but only *ahistorical* features of the agent. It is beyond the scope of this paper to address his argument.

she does. Such dispositions include being disposed to certain kinds of romantic partnerships, eat food with green curry, and daydream about getting an argument just right. In fact, most of an agent's constitutive luck falls into this banal category given the broad range of ways in which luck affects our dispositions to act. There is, however, a limited sense in which even responsibility-banal constitutive luck enables moral agency, because it provides internal boundaries that can generate reasons for action and acting for reasons is vitally important for true morally responsible action.

My explanation for why the Strawsonian generalization is a bad one is that there is an important difference between responsibility-undermining constitutive luck, on the one hand, and responsibility-enabling and responsibility-banal constitutive luck, on the other. The main difference is that the undermining kind damages or diminishes reflective self-control; the banal kind neither diminishes nor enhances reflective self-control; and the enabling kind makes possible reflective self-control. If philosophers think that the Strawsonian generalization is plausible, it is up to them to tell us why the differences that I point out between these three kinds of constitutive luck are differences that are not normatively relevant in the way that I suggest.

In the absence of that kind of explanation, I contend that this error theory dampens the intuitive attractiveness of the Responsibility Premise.²¹ And since the intuitive appeal of the

²¹ There are other error theories in the literature too. First, Benjamin Hart (2012) argues that we should understand a distinction between responsibility-undermining and -enabling constitutive luck along the lines of proper function. Second, Talbert (2015: 79) writes, “[W]e might worry that [Galen] Strawson is running together causal and moral conceptions of responsibility. Strawson identifies true moral responsibility with ultimate responsibility, but the latter seems to be primarily a causal notion: a person is ultimately responsible for some outcome only if she is the causal source of that outcome. But why should we grant that this is what true moral responsibility or true blameworthiness comes to?” Third, Wolf (2015) offers an error theory that turns on her distinction between character and mental illness. For her, character traits (including moral virtues, vices, and neither) exemplify “active intelligence” by being flexible and responsive to the agent’s perception, reason, imagination, empathy, and so on, but mental illnesses are stiff and unresponsive to them (Wolf 2015: 361-368). As a result, when a person acts from character, the person is necessarily involved in performing the action that her character explains. But then, the following rationale that might have made the Responsibility Premise appear plausible is debunked: “[O]nce one’s character is established the person’s behavior is out of her hands” (Wolf 2015: 369).

Responsibility Premise is the only reason that Strawson provides to embrace that premise, the error theory functions to diminish the plausibility of the Basic Argument itself.

In closing, I highlight a limitation for any error theory—and so for mine in particular. An error theory for a proposition is merely an explanation of its intuitive attractiveness in a way that is consistent with its being false. So, it is not an argument that the proposition is false. In my case, that there is a kind of constitutive luck that at least partially enables true morally responsible agency does not itself show that the Responsibility Premise is false. It could be the case that responsibility-enabling constitutive luck is necessary for true morally responsible action in one way and is sufficient to rule out true morally responsible action in another way. Alfred Mele (2013: 241) thinks about the luck objection to libertarianism in a structurally identical way—namely, indeterminism is necessary for true morally responsible action in one way and is sufficient to rule out true morally responsible action in another way.²² So then, my error theory is merely an explanation for the intuitive attractiveness of the Responsibility Premise; I rely on other arguments to show that the Responsibility Premise is implausible.

Concluding Thoughts

I have argued that the Basic Argument is unpersuasive and unsound. By putting the Basic Argument into the broader context of the moral luck debate, I argued that Strawson's reason for thinking that the Responsibility Premise is true is unpersuasive, because our commonsense ideas about true moral responsibility contradict the Responsibility Premise and appear to support it as well. Subsequently, I argued that common sense does *not* support the Responsibility Premise, because it is too demanding. That is, the Responsibility Premise is so demanding that *either* it is an implausible requirement for the true moral responsibility that we take ourselves to have, or it is a

²² Mele is officially agnostic about the success of the luck objection to libertarianism.

plausible requirement for a kind of true moral responsibility that we have never taken ourselves to have. Either way, the Basic Argument does not show that we lack the kind of true moral responsibility that we take ourselves to have, and so does not license a pessimistic conclusion.

Additionally, I offered an error theory for the Responsibility Premise. The Responsibility Premise gets its intuitive appeal by a generalization *from* standard excusing and exempting cases of constitutive luck that obviously undermine true moral responsibility *to* all instances of constitutive luck being such that they undermine true moral responsibility. I argued that this is an overgeneralization, because there is a relevant difference between responsibility-undermining constitutive luck, on the one hand, and responsibility-enabling and responsibility-banal kinds, on the other.

The claim that the Basic Argument is unsound connects straightforwardly to my project of defending the claim that moral luck exists and is everywhere. *Moral luck* occurs when certain kinds of factors beyond an agent's control positively affect her praiseworthiness or blameworthiness (Hartman 2017: 2). One way to *deny* that moral luck exists is to claim that luck universally undermines true moral responsibility-level control, because that claim rules out luck's ability to make a positive difference regarding the praiseworthiness and blameworthiness of the individuals in each case pair (the drivers, the judges, and the citizens). But if we can show that various ways to deny extant moral luck are implausible, then we have reason to think that moral luck exists. In Hartman (2017), I argue that Greco's (1995), Levy's (2011), and Nagel's (1979) skeptical arguments from luck are unsound.²³ If my arguments here show that Strawson's skeptical argument from luck is also unsound, then it would add more weight to the negative part of my overall argument for the

²³ In Hartman (2017), I also argue against *non*-skeptical ways to deny the existence of various kinds of moral luck as the other part of my negative case for the existence of moral luck (see also manuscript-a, manuscript-b). Additionally, I provide a positive case for thinking that various kinds of moral luck exist, and offer a global error theory for the luck-free intuition. My error theory here is local for the *constitutive* luck-free intuition.

existence of moral luck, because yet another skeptical way to deny the existence of moral luck would be implausible.²⁴

²⁴ I am grateful to Godehard Brüntrup, Daniel Haas, Luke Henderson, Andrew Khoury, Benjamin Matheson, Sean McAleer, Al Mele, Per Milam, Christian Munthe, Justin Noia, Paul Russell, Philipp Schwind, Michael Scoville, Jeremy Skrzypek, András Szigeti, Matt Talbert, and two anonymous referees for comments or conversations about the ideas in this paper. I am especially grateful to Kristin Mickelson for many conversations about the Basic Argument. I thank audiences for questions and comments at the Practical Philosophy Seminar at the University of Gothenburg, the Munich School of Philosophy Department Colloquium, the University of Zürich Department Colloquium, and the Central Division of the American Philosophical Association.

REFERENCES

- Clarke, Randolph (1997) 'On the Possibility of Rational Free Action'. *Philosophical Studies*, 88, 37-57.
- _____ (2005) 'On an Argument for the Impossibility of Moral Responsibility'. *Midwest Studies in Philosophy*, 29, 13-24.
- Coffman, E. J. (2015) *Luck: Its Nature and Significance for Human Knowledge and Agency*. New York: Palgrave Macmillan.
- Corabi, Joseph (2017) 'Two Arguments for Impossibilism and Why It isn't Impossible to Refute Them'. *Philosophia*, 45, 569-84.
- Ciurria, Michelle (2014) 'The Case of JoJo and Our Pretheoretical Intuitions: An Externalist Interpretation'. *Review of Philosophy and Psychology*, 5, 265-76.
- Cyr, Taylor C. (forthcoming) 'Moral Responsibility, Luck, and Compatibilism'. *Erkenntnis*.
- Faraci, David and David Shoemaker (2010) 'Insanity, Deep Selves, and Moral Responsibility: The Case of JoJo'. *Review of Philosophy and Psychology*, 1, 319-32.
- Faraci, David and David Shoemaker (2014) 'Huck Vs. JoJo: Moral Ignorance and the (A)Symmetry of Praise and Blame'. In Joshua Knobe, Tania Lombrozo, and Shaun Nichols (eds.), *Oxford Studies in Experimental Philosophy*, vol. 1 (Oxford: Oxford University Press), pp. 7-27.
- Fischer, John Martin (2006) 'The Cards That Are Dealt You'. *Journal of Ethics*, 10, 107-29.
- Fischer, John Martin and Mark Ravizza (1998) *Responsibility and Control: A Theory of Moral Responsibility*. Cambridge: Cambridge University Press.
- Fischer, John Martin and Neal Tognazzini (2011) 'The Physiognomy of Responsibility'. *Philosophy and Phenomenological Research*, 82, 381-417.
- Franklin, Christopher (2018) *A Minimal Libertarianism: Free Will and the Promise of Reduction*. Oxford: Oxford University Press.
- Greco, John (1995) 'A Second Paradox Concerning Responsibility and Luck'. *Metaphilosophy*, 26, 81-96.
- Hart, Benjamin (2012) *Why We Are Lucky To Be In Control: A Defense of the Compatibility of Luck and Responsibility*. Ph.D. Dissertation. St. Louis University.
- Hartman, Robert J. (2017) *In Defense of Moral Luck: Why Luck Often Affects Praiseworthiness and Blameworthiness*. New York: Routledge.
- _____ (manuscript-a) 'Against the Character Response to the Problem of Moral Luck'.
- _____ (manuscript-b) 'Moral Luck and the Unfairness of Morality'.

- _____ (manuscript-c) 'Moral Luck in Kant's Moral Philosophy'.
- Henderson, Luke (2014) 'Character Development and Heaven'. *International Journal for Philosophy of Religion*, 76, 319-30.
- Hendrickson, Noel (2007) 'Improving the Metaphysical Argument Against Free Will'. *Philosophical Papers*, 36, 271-94.
- Hurley, Susan L. (2000) 'Is Responsibility Essentially Impossible?' *Philosophical Studies*, 99, 229-68.
- Istvan, Michael Anthony Jr. (2011) 'Concerning the Resilience of Galen Strawson's Basic Argument'. *Philosophical Studies*, 155, 399-420.
- Kane, Robert (1996) *The Significance of Free Will*. Oxford: Oxford University Press.
- Kershnar, Stephen (2015) 'Moral Responsibility and Foundationalism'. *Philosophia*, 43, 381-402.
- Kment, Boris (2017) 'Free Will and Ultimate Explanation'. *Philosophical Issues*, 27, 114-30.
- Levy, Neil (2011) *Hard Luck: How Luck Undermines Freedom and Moral Responsibility*. Oxford: Oxford University Press.
- Mason, Elinor (2015) 'Moral Ignorance and Blameworthiness'. *Philosophical Studies*, 172, 3037-57.
- McKenna, Michael (2008) 'Ultimacy and Sweet Jane'. In Nick Trakakis and Daniel Cohen (eds.), *Essays on Free Will and Moral Responsibility* (Cambridge: Cambridge Scholars Publishing), pp. 186-208.
- Mele, Alfred (1995) *Autonomous Agents: From Self-Control to Autonomy*. Oxford: Oxford University Press.
- _____ (2006) *Free Will and Luck*. Oxford: Oxford University Press.
- _____ (2013) 'Moral Responsibility and the Continuation Problem'. *Philosophical Studies*, 162, 237-55.
- Mickelson, Kristin (2015) 'The Zygote Argument is Invalid: Now What?' *Philosophical Studies*, 172, 2911-29.
- _____ (2017) 'The Manipulation Argument'. In Meghan Griffith, Kevin Timpe and Neil Levy (eds.), *The Routledge Companion to Free Will* (New York: Routledge), pp. 166-78.
- Nagel, Thomas (ed.). (1979) 'Moral Luck'. In *Mortal Questions* (Cambridge: Cambridge University Press), pp. 24-38.
- _____ (1986) *The View from Nowhere*. Oxford: Oxford University Press.

- O'Connor, Timothy (2011) 'Agent Causal Theories of Freedom'. In Robert Kane (ed.), *The Oxford Handbook of Free Will*, 2 ed. (Oxford: Oxford University Press), pp. 309-28.
- Peels, Rik (2017) *Responsible Belief: A Theory in Ethics and Epistemology*. Oxford: Oxford University Press.
- Pereboom, Derk (2001) *Living without Free Will*. Cambridge: Cambridge University Press.
- Russell, Paul (2008) 'Free Will, Art, and Morality'. *The Journal of Ethics*, 12, 307-25.
- _____ (2017) *The Limits of Free Will*. Oxford: Oxford University Press.
- Shoemaker, David (2015) *Responsibility from the Margins*. Oxford: Oxford University Press.
- Smilansky, Saul (2000) *Free Will and Illusion*. Oxford: Oxford University Press.
- Strawson, Galen (1994) 'The Impossibility of Moral Responsibility'. *Philosophical Studies*, 75, 5-24.
- _____ (2000) 'The Unhelpfulness of Indeterminism'. *Philosophy and Phenomenological Research*, 60, 149-55.
- _____ (2002) 'The Bounds of Freedom'. In Robert Kane (ed.) *The Oxford Handbook of Free Will*, 1st ed. (Oxford: Oxford University Press), pp. 158-77.
- _____ (2010) [1986] *Freedom and Belief*. Revised Ed. Oxford: Oxford University Press.
- _____ (2011) 'Free Will'. In Edward Craig (ed.) *The Routledge Encyclopedia of Philosophy*. London: Routledge.
- Talbert, Matthew (2015) *Moral Responsibility: An Introduction*. Malden: Polity Press.
- Thomson, Judith Jarvis (1989) 'Morality and Bad Luck'. *Metaphilosophy*, 20, 203-21.
- Todd, Patrick (forthcoming) 'A Unified Account of the Moral Standing to Blame'. *Noûs*.
- Tucker, Chris (2007) 'Agent Causation and the Alleged Impossibility of Rational Free Action'. *Erkenntnis*, 67, 17-27.
- Wallace, R. Jay (1994) *Responsibility and the Moral Sentiments*. Cambridge: Harvard University Press.
- Waller, Bruce N. (2011) *Against Moral Responsibility*. Cambridge: MIT Press.
- Watson, Gary (1993) 'Responsibility and the Limits of Evil: Variations on a Strawsonian Theme'. In John Martin Fischer and Mark Ravizza (eds.), *Perspectives on Moral Responsibility* (Ithaca: Cornell University Press), pp. 119-50.
- Williams, Bernard (ed.). (1981) "'Moral Luck'. In *Moral Luck: Philosophical Papers 1973-1980* (Cambridge: Cambridge University Press), pp. 20-39.
- _____ (1985) *Ethics and the Limits of Philosophy*. Cambridge: Harvard University Press.

Wolf, Susan (1987) 'Sanity and the Metaphysics of Responsibility'. In Ferdinand Schoeman (ed.), *Responsibility, Character, and the Emotions* (Cambridge: Cambridge University Press), pp. 46-62.

_____ (1990) *Freedom within Reason*. Oxford: Oxford University Press.

_____ (2011) 'Blame, Italian Style'. In R. Jay Wallace, Rahul Kumar, and Samuel Freeman (eds.), *Reasons and Recognition: Essays on the Philosophy of T. M. Scanlon* (Oxford: Oxford University Press), pp. 332-47.

_____ (2015) 'Character and Responsibility'. *Journal of Philosophy*, 112, 356-72.

Zimmerman, Michael J. (1987) 'Luck and Moral Responsibility'. *Ethics*, 97, 374-86.

_____ (2002) 'Taking Luck Seriously'. *The Journal of Philosophy*, 99, 553-76.