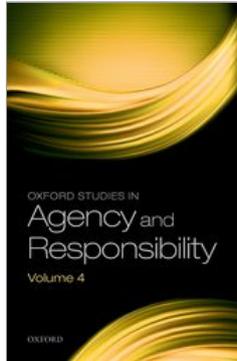


University Press Scholarship Online

Oxford Scholarship Online



Oxford Studies in Agency and Responsibility
Volume 4

David Shoemaker

Print publication date: 2017

Print ISBN-13: 9780198805601

Published to Oxford Scholarship Online: August 2017

DOI: 10.1093/oso/9780198805601.001.0001

Free Will Pessimism

Paul Russell

DOI:10.1093/oso/9780198805601.003.0006

Abstract and Keywords

The core aim of this paper is to articulate the essential features of an alternative compatibilist position, one that is responsive to sources of resistance to the compatibilist program based on considerations of fate and luck. The approach taken relies on distinguishing carefully between issues of skepticism and pessimism as they arise in this context. A compatibilism that is properly responsive to concerns about fate and luck is committed to what is described as free will pessimism, which is to be distinguished from free will skepticism. The conclusion reached is that critical compatibilism and free will pessimism should not be understood as providing a solution to the free will problem but rather as a basis for rejecting the assumptions and aspirations that lie behind it. This approach reveals not a (skeptical) *problem* waiting to be solved but a (troubling) human *predicament* that needs to be recognized and acknowledged.

Keywords: free will, moral responsibility, compatibilism, (moral) skepticism, moral luck, fatalism, the morality system, Bernard Williams, Thomas Nagel, Daniel Dennett

The will is as free as it needs to be. That does not mean, as libertarians would take it, that it is able to meet all the demands of the morality system ... Nor does it mean that it is free enough to keep the morality system in adequate business, as reconcilers

usually take it to mean. It means that if we are considering merely our freedom as agents ... we have quite enough of it to lead a significant ethical life in truthful understanding of what it involves.

Bernard Williams¹

The immediate aim of this paper is to articulate the essential features of an alternative compatibilist position, one that is responsive to sources of resistance to the compatibilist program based on considerations of fate and luck. The approach taken relies on distinguishing carefully between issues of skepticism and pessimism as they arise in this context. A compatibilism that is properly responsive to concerns about fate and luck is committed to what I describe as free will pessimism, which is to be distinguished from free will skepticism. Free will skepticism is the view that our vulnerability to conditions of fate and luck serves to discredit our view of ourselves as free and responsible agents. Free will pessimism rejects free will skepticism, since the basis of its pessimism rests with the assumption that we *are* free and responsible agents who are, nevertheless, subject to fate and luck in this aspect of our lives. According to free will pessimism, all the major parties and positions in the free will debate, including that of skepticism, are modes of evasion and distortion regarding our human predicament in respect of agency and moral life.

The argument of this paper falls into three parts. In the first section it is argued that any plausible form of compatibilism must embrace and endorse free will pessimism. Compatibilism of this kind may be described as (p.94) “critical compatibilism”, in order to contrast and distinguish it from the more orthodox forms of (optimistic and complacent) compatibilism. In the second section of the paper I offer an explanation of why it is that compatibilism has been so reluctant to embrace or accept critical compatibilism and the free will pessimism that it involves. The explanation provided turns largely on the role of what Bernard Williams has described as “the morality system” (as in the epigraph to this chapter), and its peculiar assumptions and aspirations. Finally, in the third and last section, I consider the general significance of these reflections and observations about critical compatibilism and free will pessimism and their implications for the free will problem itself. The conclusion I reach is that critical compatibilism and free will pessimism should not be understood as providing a solution to the free will problem but rather as a basis for rejecting the assumptions and aspirations that lie behind it—assumptions and aspirations that have been shared by all the major parties involved in this debate. What we have, according to the stance of free will pessimism, is not a (skeptical) *problem* waiting to be solved but a (troubling) human *predicament* that needs to be recognized and acknowledged.

I. Metaphysical Attitudes and the Free Will Problem

Before turning to the argument for critical compatibilism and free will pessimism we need to consider the general structure of the free will problem and ask, in particular, what sort of “solution” are we looking for? On the face of it, the problem seems straightforward enough. We have an image of ourselves as active agents in the world who are, in some measure, in command and control of our own destinies and the trajectory of our lives. What we do and what we become is in some relevant way up to us and depends on our own deliberations and choices. It is on the basis of possessing powers and capacities of these general kinds that we take ourselves to be moral agents who may be held accountable for our conduct and character. Various skeptical challenges may be presented to undermine and discredit this self-image. The sorts of considerations that have been advanced include reflections about God, foreknowledge, and pre-destination; science and its implications as they concern deterministic laws of nature; and so on. The solution to the free will problem under this broad canopy would be to *defeat* the skeptical challenge and provide us with some form of “vindication” or “affirmation” with respect to our self-image as free and responsible agents in the world.

(p.95) Interpreted this way, the skeptical/non-skeptical divide neatly maps onto what may be described as our “metaphysical attitudes” of optimism and pessimism. Something clearly analogous to this divide goes on with respect to the issues of the existence of God and the immortality of the soul, where the skeptical challenge is also closely associated with pessimistic worries about the human condition.² Viewed this way, these metaphysical concerns are not merely theoretical issues, the position that we take on such questions will shape our sense of the value and significance of human life itself. The issue of being disconcerted and disenchanting certainly looms before us under some forms of skeptical challenge. On this account, the relationship between our metaphysical attitudes and the parties involved in the free will dispute seems simple:

Skepticism → Pessimism

Refutation of Skepticism → Optimism

A particularly vivid example of this relationship is provided in the first chapter of Daniel Dennett’s *Elbow Room*, an influential compatibilist work that sets about the task of discrediting the “gloomleaders” of skepticism and to vindicate the “optimistic” conclusion that free will is not an illusion.³ On Dennett’s analysis, almost all the worries associated with the free will problem involve “bugbears and “bogeymen” that have been generated by misleading “intuition pumps” employed by

philosophers in the Western tradition—producing a set of groundless anxieties that require some readily available philosophical therapy for their relief.

While libertarians and compatibilists, like Dennett, may disagree about how the skeptic can be defeated they are, nevertheless, agreed that this can be done and that this serves to secure a more optimistic view of human life.⁴ However, as with the parallel cases concerning God and the immortality of the soul, not all philosophers accept that free will skepticism implies any significant or severe form of pessimism. An alternative strategy, therefore, is to defeat pessimism without refuting skepticism.⁵ From this standpoint, an optimistic solution can found without following either the libertarian or compatibilist in their non-skeptical commitments. Finally, the traditional skeptic may be unpersuaded by all these strategies and insist that not only can skepticism not be refuted, this remains a basis for pessimism about the (p.96) human predicament, an outlook which is indeed disillusioning and troubling because it discredits our self-image as free and responsible beings.⁶ In general terms, this exhausts the various available views and strategies on free will as they relate to our metaphysical attitudes and their respective grounds. In what follows it will be argued that *all* the above positions and strategies are, in different ways, guilty of *evasion* about the real nature of the human predicament and seek a “solution to the free will problem” that precludes a truthful and accurate account of what our predicament involves. This argument will begin with an argument showing that a plausible compatibilism must take the form of *critical* compatibilism and endorse free will pessimism.

II. Compatibilism, Skepticism, and Pessimism

The best way to approach this issue is by way of considering Thomas Nagel’s seminal contribution concerning the problem of “moral luck”. Nagel’s account of the problem of moral luck provides us with an especially powerful and pertinent understanding of the skeptical challenge in this sphere. The core problem in Nagel’s discussion concerns the relationship between freedom and responsibility, where this is understood in terms of the relationship between control and moral evaluation. Intuitively, Nagel argues, people can only be reasonably held responsible or subject to moral evaluation for what they have control over. However, reflection on control suggests that “ultimately nothing or almost nothing about what a person does seems to be under his control.”⁷ This observation, Nagel argues, “threatens to erode most of the moral assessments we find it natural to make.” We

can categorize the various ways in which we find that control is eroded into the following four modes of moral luck.

(1) *Constitutive luck* concerns the kind of person that we are and what our moral character is like. This includes not only our dispositions of choice but also our inclinations, capacities, and temperament. To a considerable extent these matters of character and disposition do not depend on our own prior choices or decisions (i.e., for the most part we are not “self-made-selves”).⁸

(p.97) (2) *Circumstantial luck* concerns the kinds of situations and choices that we face or encounter and must respond to. Obviously there is considerable variation in the sorts of challenges and difficulties that we may be presented with. Again, we have limited control over such factors even though they crucially influence the way in which we will be subject to moral evaluation and what we will actually be held responsible for.

(3) *Consequential luck* concerns how our actions and choices actually turn out, which includes upshots and results that may be entirely unintended and unforeseen. Nevertheless, the specific ways in which our actions and choices play out may greatly influence how we are evaluated and, indeed, whether we are praised or blamed.

(4) *Antecedent luck* concerns the final retreat to the agent’s will as a potential source of pure, untainted control. Even here, however, we may still find that antecedent conditions influence the agent’s will and, hence, in the final analysis, the agent’s own will slips away from control and is vulnerable to external, alien factors.

When we consider all these various dimensions of moral luck and the limits of control we are in danger of arriving at the conclusion that, since nothing is properly and fully under the agent’s control, there are no suitable foundations for moral evaluation or moral responsibility.

It is a notable merit of Nagel’s analysis that he stops short of endorsing the skeptical conclusion. What he aims to do is to describe a problem that appears to be recalcitrant to *any* solution. Nagel is not so much concerned to generate a skeptical argument as to put his finger on the various but related ways in which *all* parties in the free will debate—including the skeptic—may be judged guilty of evasion. (a) On the face of it, compatibilist accounts are especially vulnerable to Nagel’s analysis. Although compatibilists “leave room for ordinary conditions of responsibility”, they fail to “exclude the influence of a great deal that [a

person] has not done".⁹ That is to say, moral evaluations of conduct in these circumstances are *impure* or *tainted* because they leave scope for the influence of factors that the agent does not control (for example, in relation to her circumstances, constitution, and so on). (b) In contrast with compatibilism, libertarian accounts aim to secure a sharp, neat boundary between the active self, which decides what we do, and what is external and alien. But this (internal) self-image, Nagel argues, is eroded under reflection, whereby all that we are is consumed within the natural order of events—leading eventually to the disappearance of the active agent, who gets swallowed up in the (p.98) causal flow of nature. (c) Finally, despite the apparent slide into skepticism, Nagel firmly resists any easy solution of this kind because it is, as he describes our experience, impossible for us to entirely abandon or jettison our "internal" sense of self as a responsible agent, or to refrain from our tendency to extend this view of ourselves to others.¹⁰ According to Nagel's analysis, "solutions" in any of these directions fail to fully acknowledge that we are simultaneously pulled in *two opposing directions* that we cannot reconcile. The standard strategies that Nagel considers each try to collapse the problem by emphasizing one side of the dilemma rather than the other—but to do this is mere evasion rather than solution to the problem that we encounter.

One feature of Nagel's analysis that deserves particular attention, and is especially relevant for understanding the approach taken by critical compatibilism, is what he takes to be the core requirement for any adequate attempt to preserve free and responsible agency. What is crucial, of this account, is that the active self—the free, responsible agent—must be *insulated* from the influence of fate and luck. Although Kantians and libertarians understand this general requirement in more specific terms relating to securing some form of sourcehood or ultimate agency this is, nevertheless, a requirement that *all* parties in the free will debate accept under some interpretation. It is, moreover, a key assumption that that does much to shape the entire "free will problem" and the debate that surrounds it. If this general requirement cannot be met, it is agreed by all parties, then our self-image as free, responsible agents will be compromised and will collapse.

The general requirement described above plays a key role in the core incompatibilist argument against all compatibilist strategies and proposals. Let us call this incompatibilist argument the *Basic Exclusion Argument* (BEA):

1. There is a set of conditions φ (under some contested interpretation) such that an agent is free and responsible for an action or set of actions when these conditions are satisfied.

2. There is another set of conditions β (under some contested interpretation) such that an agent's action or set of actions are subject to fate and luck when those conditions are satisfied.

*3. Any action (or set of actions) that satisfy φ cannot be such that it also satisfies β . That is to say, if an action X satisfies φ it cannot also be subject to β . <Exclusion Premise (EP)>.

4. Any and all compatibilist interpretations of φ are such that they may be satisfied and still be subject to β (i.e., compatibilist conditions φ^* do not support or satisfy EP#3 above).

(p.99) 5. It follows that we must reject any and all compatibilist interpretations φ^* , as they are inadequate as judged by a standard that compatibilists do not and cannot reject (EP).

Libertarians believe that their own interpretations of conditions φ can satisfy EP and avoid the skeptical conclusion (although this requires the truth of indeterminism). Skeptics maintain that there is no available set of conditions φ that serve to satisfy EP and, hence, the skeptical conclusion goes through either way. In what follows I want to focus on the compatibilist response to BEA and the stance compatibilists take with respect to EP.

Proponents of BEA are entirely justified in claiming that compatibilists have consistently adhered to EP and aimed to satisfy it. What compatibilists have denied is premise #4, the claim that compatibilism fails to satisfy the standard set by EP (premise #3). Let us consider the classical compatibilist argument that is launched against premise #4, an argument aiming to show that agents who satisfy suitably interpreted compatibilist conditions (φ^*) are not subject to fate and luck (i.e., conditions β). The core feature of this argument is that the incompatibilist claim (premise #4) relies on a basic confusion between fatalism and determinism. More specifically, it is argued that if we properly interpret conditions β (i.e., β^*) then premise #4 is groundless. *Fatalism* is the doctrine that all our deliberations and actions are *causally ineffective* and make no difference to the course of events. Nothing about the thesis of determinism implies that this is the universal condition. Dennett provides a particularly vivid example of this contrast:

Consider the man who has thrown himself off the Golden Gate Bridge and who thinks to himself as he plummets, “I wonder if this is really such a good idea.” Deliberation has indeed become impotent for this man ...¹¹

While conditions of “local fatalism” of this sort may occur, and deliberation and action may sometimes be futile, circumstances of this kind are “abnormal” in a deterministic world, where deliberation is generally effective. Let us call this “contributory fatalism”, where this is understood to involve the *causal impotence* of the agent with respect to some outcome or upshot.

The critical compatibilist response to this line of argument, which aims at defending compatibilism and defeating BEA, tracks incompatibilist concerns. More specifically, the critical compatibilist agrees with the incompatibilist that appealing to the distinction between determinism and contributory fatalism is a shallow and evasive understanding of incompatibilist concerns. The relevant issue is not about the causal influence *of* the agent but rather the causal influences *on* the agent. On the assumption of (p.100) determinism, however complex the mechanisms or capacities involved, the ultimate source or origin of conduct and character is external to the agent and not within the agent’s control or influence. Fatalistic concerns of this kind, which we may term “origination fatalism”, cannot simply be set aside or ignored on the basis of considerations relating to contributory fatalism.

What these observations reveal is that, within the structure of compatibilist commitments, whatever specific form they may take, we inevitably encounter *limits* to control and the way it is actually exercised and occasioned. Neither second-order (hierarchical) capacities nor reason-responsive abilities will enable us to evade this implication.¹² What this reveals is the fact of our finitude and contingencies—these being circumstances under which all human agents inescapably must operate. While libertarians may aspire to escape limitations of this kind (for example, by postulating “unconditioned conditions”, “contra-causal freedom”, or similarly motivated forms of metaphysical apparatus of this general kind), compatibilists reject all such aspirations as illusory. Having said this, compatibilists are in no position to refuse to acknowledge the force of fatalistic concern with respect to origination issues.¹³ It is at this juncture where critical compatibilists diverge from their complacent (optimistic) compatibilist brethren. At the same time, critical compatibilists also diverge from incompatibilists—libertarians and skeptics alike—in rejecting the view that considerations of this kind, relating to origination and the limits of control, license skepticism

about freedom and moral responsibility. The capacities described by compatibilists (i.e., as identified by ϕ^* —reason-responsiveness, etc.) are, they maintain, robust and substantial enough to serve as a secure foundation for our attitudes and practices associated with moral responsibility.

At this point, the incompatibilist is sure to raise the following objection. While critical compatibilists are correct in acknowledging the force of fatalistic concern relating to origination and the limits of control, as generated on compatibilist models, the attempt to separate issues of fate and responsibility in the manner proposed cannot be acceptable. More specifically, for reasons highlighted in Nagel's discussion, the presence of conditions of origination fate bring with them worries about moral luck; that is, worries relating to agents being subject to moral evaluation in ways that are sensitive to factors that they do not control. This remains the core incompatibilist objection to the compatibilist project and concessions about fate do not address or settle *this* difference. Granted that it is intuitively unjust to hold agents responsible for aspects of their conduct and character that they do not (p.101) control (as per the exclusion thesis), conditions of freedom and responsibility cannot be sustained in circumstances where an agent is subject to fate and luck along the lines described. From this perspective, fate and luck come together, and where such conditions hold, free and responsible agency is eroded into nothing.

The usual compatibilist reply to this, as found prominently in Dennett's *Elbow Room*, is to try and deflate the luck objection. It is Dennett's basic contention, consistent with much contemporary compatibilist thinking, that human agents are "not just lucky", we are "skilled self-controllers"—this being a theme to which Dennett devotes much of his book.¹⁴ Once again, this general line of reply seems not to engage with the real force or basis of incompatibilist concern. Incompatibilists recognize, of course, that compatibilist accounts of self-control and reason-responsiveness do not leave us "merely lucky" or unskilled, unable to enhance our abilities and talents. The point is, rather, that the specific capacities we may have, the way we actually exercise them, and the occasions we are provided for employing them, all depend, given deterministic assumptions, on external factors and conditions that no agent ultimately has control over. In other words, from an incompatibilist point of view, even on the most generous and robust interpretation of compatibilist powers of rational self-control, we still face limits of control over: (1) the *acquisition* of the relevant capacities involved; (2) the way these capacities are actually *exercised* in given circumstances; and (3) the *occasions* in which these capacities must be

employed or exercised (i.e., the sorts of moral challenges we may face or be presented with). In all these cases, what we do and will be held accountable for depends on these external or alien factors. From this perspective, moral life becomes hopelessly vulnerable to luck or the limits of control, which is not permitted by the exclusion thesis and is unacceptable to all those who endorse it. Interpreted this way, the incompatibilist assessment of critical compatibilism is that it is an inherently *unstable* effort to respond to incompatibilist concerns about fate and luck, since any effort to acknowledge and accommodate those concerns, along the lines proposed, must discredit the compatibilist component of its commitments. More specifically, the attempt made by critical compatibilists to acknowledge and accommodate these concerns relating to fate and luck plainly violate EP (premise #3), which has hitherto been accepted by *all* parties in the debate.

It should be evident that, whatever the merits of the incompatibilist rejoinder described above, the critical compatibilist reply to BEA is very different to that pursued by orthodox compatibilism. Critical compatibilists (p.102) accept premise #4—they *agree* that compatibilist conditions φ^* may be fully satisfied and the agent or actions concerned still subject to relevant forms of fatalism and luck. Critical compatibilists deny, nevertheless, the skeptical conclusion because they deny EP or premise #3 (contrary to their orthodox brethren). It is the burden of the argument, so far, that a sensible, credible compatibilism is constrained by the nature and character of its own commitments to take the form of *critical* compatibilism and thus must deny EP. Failing this, compatibilism is plainly guilty of evasion and superficiality on the matters of fate and luck, just as its incompatibilist critics have suggested.

Clearly, then, the point that needs emphasis for our present purposes, is that *any plausible* form of compatibilism must recognize and acknowledge the influence of fate and luck on the manner and context in which our capacities of rational self-control operate. In consequence of this, it must reject the EP and allow that conditions of free and responsible agency may coincide with the presence of conditions of fate and luck, understood in terms of external factors beyond our control that directly influence how our capacity of self-control is actually exercised. Nagel describes circumstances of this kind in these terms:

A person can be morally responsible only for what he does: but what he does results from a great deal that he does not do;

therefore, he is not morally responsible for what he is and is not morally responsible for.¹⁵

For Nagel, embracing this outlook would leave us “morally at the mercy of fate” and so must be rejected. My argument, so far, has been that this is exactly what any plausible compatibilism must commit us to, consistent with agents possessing and exercising (robust) capacities for rational self-control.

There is another important feature of critical compatibilism that flows from the rejection of EP that needs further, independent articulation, and description. This feature concerns the metaphysical attitudes that this stance naturally licenses or occasions. In circumstances where EP is not satisfied, we have (deep) reasons for being “troubled” or “disconcerted” by our predicament as this relates to human ethical life and moral agency. Even if we are “fortunate” in the particular ethical trajectory our lives may take, there is no basis (as incompatibilists rightly insist) for an easy optimism when fate and luck intrude into our ethical lives and the way we may exercise our moral agency. These observations and reflections may and should occasion a sense of “disenchantment” about our predicament, and to this extent this will license and occasion a significant sense of pessimism (on analogy with related metaphysical issues and the attitudes that they may occasion).

(p.103) However, the crucial point in relation to critical compatibilism, is that a pessimism of this nature is not rooted or grounded in skepticism about free will and moral responsibility. On the contrary, it presupposes that we *reject* any skepticism of this kind, since the form of pessimism that is occasioned depends on viewing ourselves and others as agents who are free and responsible but, nevertheless, subject to fate and luck in the exercise and operation of our moral capacities. Let us call this stance or metaphysical attitude *free will pessimism*. I will return in the next section of this paper to say more about the nature and grounds of free will pessimism. For now, however, suffice it to note that even if we reject compatibilism (for example, because we retain a commitment to EP, as incompatibilists certainly will do) it is still crucial to recognize the significance of these findings both as they relate to critical compatibilism and the free will pessimism that flows from it.

III. Compatibilism and “The Morality System”

The question I now want to turn to is why have *compatibilists* been so reluctant to embrace critical compatibilism and free will pessimism? Incompatibilists maintain that compatibilists conditions φ^* are such that they do not exclude conditions β . Whereas orthodox compatibilists attempt to refute this premise (#4) critical compatibilists maintain that compatibilists should accept or recognize the truth of premise #4 and should instead reject EP (#3). What is it about EP that orthodox compatibilists find so difficult to abandon? There are, I suggest, two considerations that run deep in orthodox compatibilist thinking that account for this resistance to jettisoning EP. The first concerns the relation between the exclusion thesis and “the morality system” and the second, related to the first, concerns the question of optimism.

(1) The exclusion thesis may be understood as an essential feature of what Bernard Williams calls “the morality system”.¹⁶ The morality system, as Williams describes it, places particularly heavy emphasis on the (peculiar) concept of obligation, along with the closely concepts of blame and voluntariness. Moral responsibility”, as “the morality system” understands it, is taken to be primarily a matter of rational agents voluntarily violating their obligations and, thereby, being liable to blame and retribution. A further closely related feature of the morality system is that insists that moral responsibility, interpreted in these (narrow) terms, must somehow be capable of “transcending luck”, providing a purity that (p.104) only genuine (rational) agency of some kind makes possible.¹⁷ Within this framework, the aspirations of libertarianism and its commitment to EP is entirely intelligible. Although orthodox compatibilists resist the aspirations of libertarians, and its efforts to secure some form of absolute or ultimate agency, they remain committed to the particular conception of responsibility encouraged by the morality system and believe that it can be satisfied within compatibilist constraints.¹⁸ In contrast with this, critical compatibilism involves *rejecting* core features of “the morality system”, including its particular conception of moral responsibility (all this being something, if Williams is right, that we have good reason to do in any case). Although abandoning EP certainly makes it impossible to salvage the particular conception of freedom and responsibility promoted by the morality system, this is not to be confused with skepticism about freedom and responsibility *tout court*. On the contrary, while proponents of the morality system tend to present the situation this way, it is generally recognized, even by the proponents of the morality system themselves, that the narrow conception of moral responsibility constructed around the assumptions of the morality system is one that is both “local” (modern, Western) and is widely contested—including within our own modern, Western ethical community.¹⁹

(2) There is, as already mentioned, another consideration, closely related to the first, that is also very significant in this context. The aspiration to *optimism*, in particular to tell a comforting story about the human predicament in respect of moral agency, is one that runs deep in the morality system. This deep resistance to a disturbing or troubling view of human ethical life, one where the excise and operation of our moral and rational capacities depends in large measure on factors that are not controlled or governed by those same capacities and powers, is one that is not only shared by libertarians and compatibilists but that also motivates the skeptics. All of these parties, in their various ways, hold on to EP and the form of optimism that it insists on (i.e., that human ethical life does not function or operate in violation of the constraints that EP imposes upon it). Put in other terms, the form of optimism that EP insists on is one that rejects the very *possibility* of free will pessimism, much less accepts it as the *truth* about our human predicament. It is within this philosophical fabric, as encouraged by the (p.105) forms of theorizing associated with the morality system, that (orthodox) compatibilist resistance to abandoning EP should be understood and appreciated. Clearly if we allow that free and responsible action may nevertheless be infused with conditions of fate and luck, we must also abandon any form of unqualified optimism—in particular, the hyper-optimism that compatibilists such as Dennett endeavor to project.²⁰

Critical compatibilism endorses no form of easy, complacent, or unmixed optimism on this subject. On the contrary, in giving *weight* to the limits of control, and circumstances of finitude and contingency in the sphere of human agency, critical compatibilism suggests *a particular understanding of pessimistic concern*—namely “free will pessimism” (as opposed to skepticism about freedom and responsibility). We might describe this stance as one as that recognizes or acknowledges that conditions of freedom and responsibility do not *elude* those of fate and luck but rather *confront* fate and luck and that these conditions are, indeed, meshed and entangled together. All theories and interpretations that deny this are, from this perspective, guilty of various modes of evasion that involve some effort of one kind or another to satisfy EP and the forms of optimism associated with it. This particular aspiration is something that critical compatibilists maintain we must abandon, not only because it generates insoluble philosophical perplexities but, more importantly, because it

misrepresents the (difficult and troubling) *truth* about our circumstances as human agents.

One way to resist free will pessimism is to reject the suggestion that if we abandon EP then some form of pessimism must follow from this. Perhaps, critics may argue, no relevant metaphysical attitude of this (negative) kind needs to be generated by reflections of this nature. With respect to this matter a few general observations about the nature of the pessimism involved in free will pessimism are called for. (a) We can, as we have done, distinguish critical compatibilism from its orthodox counterpart simply with reference to the disagreement about EP (and how to respond to BEA). To this extent it may be argued that no specific affective or attitudinal element is *essential* to the core distinction between these two compatibilist stances as presented. (b) Moreover, we may also concede that whatever pessimistic features may naturally accompany the rejection of EP and the associated apparatus of the morality system, there is still much to *welcome* about this shift away from these commitments and aspirations (for example, we are better off without pernicious and destructive forms of retributivism that (p.106) are grounded in these views).²¹ For this reason critical compatibilism may involve elements of *both* optimism and pessimism. There is no reason to suppose that the relevant metaphysical response here must be unmixed or uniform in nature—a finding that should not surprise us since the same is obviously true with respect to other, similar metaphysical issues and reflections (for example, concerning the existence of God, immortality, and so on). (c) The *degree* of pessimistic affect may vary depending on both our historical and cultural circumstances. For those who find themselves deep inside the morality system and its assumptions and aspirations, the sense of being troubled and disturbed by the thought of abandoning EP may well be amplified in proportion to the depth of their existing set of commitments. We may allow, therefore, that the sense of pessimism is likely to recede or dissipate over time as *we* (moderns, Westerners) withdraw from the morality system. All these concessions are consistent with the claim that critical compatibilism still licenses a distinct form of pessimism.

We may be pressed further here and asked to say more about the character of free will pessimism and how it relates to the critical compatibilism and the rejection of EP. In describing any form of pessimism we need to make reference to two central features: the grounds or basis of the pessimism and the quality or affective aspects that this may involve or imply. We have already been careful to emphasize that the grounds of free will pessimism presuppose that skepticism about freedom and moral responsibility is mistaken. Plainly

the particular source of pessimism must be traced back to the view that free and responsible agents are, nevertheless, subject to significant forms of fate and luck (contrary to the requirements of EP). In general, pessimism is called forth or occasioned when something that we value is threatened or discredited. In this case, the significant feature of our condition that is discredited and undermined concerns our (deep) assumption and hope that there is some sort of “harmony” between our ethical life and the order of the world.²² It is, of course, a central theme of Williams’ work to argue that the ancient Greeks, unlike *we* (Western) moderns, had no such expectations about the world and how it is related to human ethical life. The important point here is that the move away from the morality system is not *cost-free*—whatever gains and advantages it may involve in other respects (including that of truthfulness).

Some similar and related observations are also in order with respect to the quality or affective aspect of free will pessimism. Pessimism may take (p.107) various affective forms and degrees, ranging from intense despair and grief to a milder sense of being disconcerted or disenchanted with the state of things. Consider, for example, Pascal’s attitude to the supposition that there was no future state, which took the form of an extreme and severe pessimism that we may describe as excessive and immoderate. At the same time, we may also regard the Epicurean response of complete equanimity in face of death as being too shallow and lacking appropriate sensitivity to features of human life that should be recognized as troubling and difficult. It is within a matrix of this general kind that we should understand the stance of the free will pessimist. The response that is brought forth on the basis of reflections about our predicament as free and responsible agents who are, nevertheless, subject to conditions of fate and luck in the very exercise of our agency, is one that falls *between* the extreme of Pascalian despair and Epicurean calm and complacency. The right response, one that is duly sensitive to the features of our predicament, is that of a moderate but engaged sense of being disenchanted or disconcerted. Orthodox compatibilism is in this sense not only philosophically *evasive* with respect to the issues that it confronts, the philosophical evasions manifest themselves in forms of affective *shallowness or superficiality*—a failure to care about (problematic) features of our predicament that are or should be significant to us.

It is arguable that the preceding analysis of the source and nature of orthodox compatibilist resistance to critical compatibilism and free will pessimism still fails to identify the core concern. More specifically, it may be argued that orthodox compatibilist resistance to critical

compatibilism stands at no great distance from incompatibilist resistance to compatibilism in general. The relevant issue here has to do with EP and the force of BEA itself. Orthodox compatibilists reject BEA because they reject the claim that compatibilist conditions may be satisfied when EP is violated (i.e., they deny premise #4). Their unwillingness to reject EP is rooted in the concern that if EP is not satisfied, or at least respected, then *morality itself would be unfair*. It would be unfair to endorse and apply conditions of freedom and responsibility that leave agents vulnerable to fate and luck. The thought here is one that supports much of Nagel's entire analysis of the problem of moral luck and the associated fabric of the free will problem. If fate and luck are infused into the very exercise and operation of our moral capacities as agents then all moral evaluation is tainted and impure. Morality, as it were, requires fairness *all the way down*—and that is what EP, it is claimed, secures for us.²³

(p.108) Clearly critical compatibilism rejects this view of things and the assumptions and aspirations that must accompany it (and which lead on to the intractable nature of the free will problem itself). For critical compatibilism, which holds that satisfying EP is not a necessary condition of sustaining human freedom and moral responsibility, the relevant standard of fairness of the attitudes and practices associated with moral responsibility are provided and adequately secured by the satisfaction of the set of compatibilist conditions identified by φ^* . The issue of fairness is *internal* to that fabric, which is itself liable to be modified, amended, and corrected over time.²⁴ Any demand for fairness beyond this—some mode of *absolute* fairness—is liable to simply collapse under its own weight and results, ironically, in skepticism. This, at least, is where critical compatibilists (must) stand on this issue. Insisting on these observations, as they relate to human ethical life, is not, however, one that secures or encourages any simple, unmixed form of optimism. On the contrary, when we consider the various possible ethical trajectories human lives may take—especially those that are ethically unfortunate and problematic—we are confronted with an awareness of our ethical fragility and vulnerability, consistent with conceiving of ourselves as free, responsible ethical agents.²⁵ The very fact that we are and should be troubled by reflections of this kind is evidence that there is something deeply wrong with the assumptions of the morality system and with EP in particular. It is only on the basis of an understanding and appreciation of (the possibility of) free will pessimism that we can make any adequate sense of our response to these cases and the particular way in which we find them disturbing (i.e., that we recognize that agents are both free and responsible and subject to conditions of fate and luck, contrary to EP). Much of the

resistance to critical compatibilism that comes from within compatibilism itself may be accounted for as rooted in the discomfort that is felt once we acknowledge that EP does *not* govern human ethical life. This is, no doubt, a hard truth about our ethical lives that many—especially for those who occupy a position well inside the morality system—will find difficult to accept. Indeed, many will find free will pessimism to be harder to accept—and much more disturbing—than any form of free will skepticism, since skepticism, at least, provides the consolation of ensuring that free, responsible agents are not subject to fate and luck (as per EP). There is an (p.109) important sense, therefore, in which free will pessimism may be found to be much more disturbing than any form of free will skepticism—and that is especially true, of course, for those who do not find skepticism disturbing in any case.

IV. Incompatibilist Evasions and Free Will Pessimism

It has been argued, so far, that a plausible compatibilism must take the form of critical compatibilism and embrace free will pessimism. It has also been argued that the best explanation for why compatibilists have hitherto been reluctant to accept (or even consider) this view is that compatibilists have generally remained committed to the assumptions and aspirations of the morality system, including EP, which is an essential feature of it. None of this, however, in itself, shows that we should accept critical compatibilism and acknowledge that free will pessimism is the truth about the human predicament as it concerns moral agency. On the contrary, incompatibilists may *welcome* the conclusions we have arrived at with respect to the implications of compatibilism. The reason for this is that incompatibilists will suppose that the argument advanced so far, concerning compatibilism and free will pessimism serves not so much as an effective *defense* of a (modified or refined) compatibilism as a *reductio* of the whole compatibilist project. An approach of this kind, the incompatibilist may say, is not so much a case of “biting the bullet” as of shooting oneself in the head. Since critical compatibilism concedes that compatibilism implies free will pessimism and necessarily violates EP, the correct conclusion to draw from all this is that we should *reject compatibilism*.

The argument that follows continues from where BEA, and the above line of criticism, leaves off: namely, with the claim that we must reject compatibilism because it fails the standard of EP. The right place to begin, therefore, is with the alternatives that incompatibilism offers us. From the perspective of critical compatibilism, however, none of the familiar incompatibilist strategies can survive critical scrutiny. They fail, in particular, the very standard of EP that incompatibilists appeal to in framing BEA and are no less guilty of their own distinct forms of evasion (i.e., no less guilty than orthodox compatibilism). For our present purposes, which is to identify and present the core structure of the critical compatibilist argument leading to the conclusion that we should reject EP and endorse free will pessimism, what is required is, first, a reminder of the relevant reasons for rejecting incompatibilism and the various modes of evasion that this may involve, and, second, an interpretation of the implications of this for our overall (p.110) understanding of the free will problem itself. Having explained in general terms why the retreat back to incompatibilism is not a viable or credible option, we face a clear choice between EP (and the intractable free will problem that it generates) and free will pessimism. We have every reason, I claim, to opt for free will pessimism. I will call the complete argument, extending beyond BEA, *The Extended Argument to Free Will Pessimism* (FWPA).

The structure of this argument, following on from BEA and leading to the conclusion that we should reject EP and embrace free will pessimism, takes this form:

1.-5. [BEA (#1- #5)]

6. If not compatibilism (i.e., because it fails EP, as identified by BEA), then incompatibilism.

7. If incompatibilism, then either libertarianism or skepticism.

8. If libertarianism, then either (a) agent-causal or (b) event-causal libertarianism.

9. Both agent-causal and event-causal forms of libertarianism are guilty of evasion with respect to the issues raised by EP. The first appeals to unintelligible or incoherent forms of (“panicky”) metaphysics; and the second is vulnerable to modes of luck of the kind that are proscribed by EP.

10. Given that both compatibilism (BEA) and libertarianism fail (FWPA, #6-#9), skepticism must follow. Although skepticism does not aim to *satisfy* EP (in contrast with orthodox compatibilism and libertarianism), it still *respects* EP. Since EP cannot be satisfied—that is, there are *no* conditions φ that satisfy EP—a skeptical conclusion must follow (one that is generally taken to license its own form of pessimism).

11. Skepticism is either (a) global or (b) local. Local skepticism targets modes of responsibility that are specifically encouraged by “the morality system” and that aim to satisfy EP. Global skepticism maintains that there are *no* credible accounts of conditions φ on the ground that any acceptable account of conditions φ must satisfy EP (i.e., local skepticism and global skepticism are identical).

12. Local forms of skepticism are not a threat to free will pessimism or critical compatibilism, since it also endorses local skepticism of this kind (i.e., accepts and acknowledges that EP cannot be satisfied). Global forms of skepticism are a threat to free will pessimism and critical compatibilism since they endorse the claim that any acceptable account of conditions φ must satisfy EP, which rules out the very possibility of free will pessimism.

13. Global skepticism involves “bad faith” and is itself a form of evasion with respect to (the possibility of) free will pessimism. More specifically, (p.111) global skepticism involves denying the significance of moral capacities and abilities that human beings evidently possess and

exercise. Skeptics generally begin by advancing global claims and then retreat back to local skepticism, manifesting their own discomfort with the evasions of global skepticism.

14. When we apply EP, and aim to satisfy or simply respect its demands, we find that it generates a range of unacceptable and unconvincing forms of evasion.

15. The free will problem structured around BEA and EP is, as Nagel's original analysis suggests, *intractable*. The root source of this impasse rests with "the morality system" and its commitment to EP, which denies the very possibility of free will pessimism (and thus critical compatibilism).

16. We are faced with a fundamental choice between EP and free will pessimism. In light of the preceding analysis, we have every reason to reject EP and embrace free will pessimism.

To show that FWPA is entirely sound we would need to defend each of the premises relating to the challenges facing incompatibilism (#8-#13). As I have indicated, there is no reason, in the present context, to repeat and rehearse all the relevant arguments and considerations in this context. It will suffice to provide a general overview of the core concerns and objections that serve to justify and support the premises concerned. It is reasonably obvious that if we reject compatibilism because it fails EP then we must consider one or other of the incompatibilist alternatives, libertarianism or skepticism (as per #6 and #7). Let us begin with the problems facing libertarianism.²⁶ The libertarian believes that there is some available account of conditions ϕ (i.e., conditions $\phi\#$) that satisfy the requirement of EP but that this depends on the falsity of determinism. In order to satisfy EP, with a view to overcoming the limits of control in a way that provides for genuine ultimacy (and "self-creation"), libertarianism in its classical form introduced what even its most prominent exponents have described as an "odd" set of metaphysical commitments.²⁷ The constructive metaphysical foundations of this theory, beyond the requirement of indeterminism, rests with the suggestion that free and responsible agents have active powers "which some would attribute only to God: each of us, when we act, is a prime mover unmoved."²⁸ The theory involved here is that of agent-causation and it has been extensively and effectively criticized not only by compatibilists but also (p.112) by some leading representatives of contemporary libertarianism. It is argued by Robert Kane, for example, that agent-causation theories are incapable of being "reconciled with modern scientific views of human beings" and that they fail to provide any

intelligible account of how genuine ultimacy is possible.²⁹ For this reason, Kane and others have defended event-causal libertarianism, a metaphysically more modest or “softer” theory.³⁰ What is crucial to event-causal theories is the suggestion that free action of the kind required for moral responsibility can be secured by a theory of agency that allows that actions may be caused by (plural) available reasons or motivations that, nevertheless, do not necessitate or determine their occurrence. Event-causal theories of this kind are, however, vulnerable to the “luck objection”—an objection rooted in the requirements of EP (as presented in BEA). The fundamental problem here, as pressed by other libertarians as well as by skeptics or hard incompatibilists, is that agents may satisfy event-causal conditions of the kind described and still be *no more* in control of their conduct than compatibilist agents. If this is the case then the condition described by event-causal theories fall short of the sort of (total) control that ultimacy or genuine sourcehood requires.³¹ Clearly, then, given that libertarianism is found wanting, this forces a further retreat back to skepticism.

The skeptic maintains that EP cannot be satisfied by any proposed set of conditions φ concerning freedom and responsibility. They all fail, in various ways, the standard imposed by EP and BEA. However, according to the skeptic, EP must still be *respected*, even if it cannot be satisfied. To this extent, skepticism is itself an expression or manifestation of “the morality system” and its core assumptions and aspirations. It expresses, as it were, both the disappointment that human agents fail the standard set by EP and the continued optimism that this standard, nevertheless, remains in place, even if human agents fail to live up to it. In order to support this stance, the skeptic must rely on the key claim that any conception of freedom and responsibility that does not satisfy EP cannot be the “true” or “genuine” conception that we are (or should be) concerned with. That is to say, the skeptical position rests on a narrow and restricted conception of (true) freedom and responsibility that satisfies the preferred requirements of the morality system and its commitment to EP. Any alternative account must, (p.113) therefore, be judged “shallow” or “superficial,” such as we find in the case of “the economy of threats account.”³² From the perspective of critical compatibilism, the skeptical position is itself just another mode of evasion generated by the morality system, one that similarly seeks to rule out the very possibility of free will pessimism (as denied by any continued commitment to EP).

Critical compatibilists will certainly agree, with the skeptic, that the particular conception of freedom and responsibility that the morality system aims to secure should be rejected and discarded. What it denies,

however, is that this licenses any unqualified form of global skepticism (i.e., as based on the general application of BEA to all proposed accounts of conditions φ). According to this view of things, we need to distinguish carefully between two distinct skeptical views; (a) a qualified or local skepticism as it concerns our Western, modern conception as encouraged by the morality system, and (b) an unqualified or global skepticism that extends to any and all proposed accounts of conditions φ . While it may be true that the narrower, local conception fails (for reasons pointed out by BEA), it does not follow from this that all proposed concepts must fail—unless, of course, we simply assume that EP holds or applies in respect of all understandings of freedom and responsibility. If this were the case, then local skepticism would directly lead to the global skeptical conclusion.

Critical compatibilists, as we have noted, are themselves committed to local skepticism as it has been interpreted above, a concession that is entirely consistent with their own compatibilist commitments (unlike orthodox compatibilists who aim to satisfy the demands of EP and the morality system from *within* their compatibilist commitments).³³ Critical compatibilists can also accept that some proposed versions of compatibilist conditions φ^* —such as “the economy of threats account”—are inadequate or insufficiently robust for the purpose of providing a substantial and credible theory of freedom and responsibility. None of this leaves them in the situation of having to accept an unqualified or global form of skepticism about freedom and responsibility, which is the only form of skepticism that is directly problematic for the critical compatibilist. The modes of freedom and responsibility grounded in the various robust, complex capacities identified and explained by (critical) compatibilism are more than adequate to the task of grounding and justifying attitudes and practices that are recognizably part of the fabric or moral and ethical life more broadly conceived (i.e., more broadly conceived than the morality system would suggest). Even skeptics generally concede this (p.114) point and attempt to mask it by initially advancing strong global skeptical arguments and then retreating back to (the more modest and more plausible) qualified skeptical conclusion that is limited to the local conceptions of freedom and responsibility that are encouraged by the morality system.³⁴

The general objection that critical compatibilism raises against (global) skepticism is that it is a form of “bad faith”. Compatibilist models of responsible agency, constructed along the lines of rational self-control and reason-responsiveness, while they may not satisfy EP, are plainly far more robust and sophisticated in accounting for a wide range of

distinctions and discriminations in this sphere than the economy of threats account. Any attempt to simply dismiss the powers and abilities described, and treat them as irrelevant to human ethical and social life, with no bearing on our reactive attitudes and retributive dispositions and practices, has more than a taint of “bad faith” about it. When we take this road, the humane goal of unmasking the distortions of the morality system and its preferred account of freedom and responsibility gets turned on its head and becomes the dehumanizing hypothesis that there are no *real* or *genuine* free and responsible agents in the world. Any skepticism of this unqualified kind must inevitably face the compelling objection that it constitutes just another mode of evasion, as encouraged by the morality system itself and its attachment to EP.³⁵

Where do these critical reflections about the prospects of incompatibilism leave us? If we follow the argument leading from BEA and the rejection of compatibilism through to the difficulties facing the various forms of (p.115) incompatibilism, we arrive back in the situation that Nagel diagnosed so forcefully in “Moral Luck”, which is the intractable nature of the free will problem. All the familiar approaches and strategies are, as Nagel suggests, guilty of evasion of one kind or another. Some attempt to solve the problem by proposing conditions φ that will satisfy EP by postulating extravagant and unintelligible metaphysical apparatus. Others seek to conceal or deny the ways in which their preferred theories remain vulnerable to concerns relating to fate and luck. Skeptics, the last representatives of the morality system, as it collapses under the weight of its own assumptions and aspirations, denies human freedom and moral responsibility altogether, rather than abandon its commitment to EP and the forms of optimism that it aims to secure. These modes of evasion are all encouraged by the morality system and the forms of theorizing that it generates. This leaves all those who accept EP, along with the free will problem constructed around it, trapped in a philosophical labyrinth that offers no way out.

According to critical compatibilism the insoluble nature of this philosophical conundrum serves as strong evidence that it is based on faulty assumptions and aspirations. The root source of this impasse rests with the morality system and its commitment to EP, which denies the very possibility of free will pessimism and critical compatibilism. We are, then, faced with a clear choice between EP and the morality system on one side, and free will pessimism and critical compatibilism on the other side. How do we decide between them? Faced with this choice between EP and free will pessimism we have every reason to opt for free will pessimism. One reason for this is that it allows us to set

aside the (intractable) free will problem and turn our attention, instead, to the distinct questions arising from the need to provide a credible account of conditions φ —unencumbered by the (faulty) requirements of EP. More importantly, however, the right basis for rejecting EP in favor of free will pessimism is that we find, on critical reflection, that free will pessimism is the most *truthful* account of our human predicament in respect of these matters. Unlike the forms of theorizing associated with the morality system, free will pessimism involves no evasions or metaphysical fabrications. It is free will pessimism, rather than the philosophical theories constructed around EP, that most accurately and adequately capture our discomfort when we reflect on troubling cases where free and responsible agents are, nevertheless, plainly entangled in circumstances of fate and luck. Given all this, we must reject EP and accept that free will pessimism is the truth about the human predicament in respect of these matters.

Granted that we should reject EP and accept free will pessimism, there is no principled basis for rejecting critical compatibilism (contrary to BEA). That is to say, the fact that compatibilist conditions φ^* fail EP, and do not satisfy the assumptions and aspiration it embodies, is not, in itself, an (p.116) acceptable basis for rejecting these conditions. The conditions φ^* proposed may, of course, be judged *more or less* adequate in terms of their descriptive accuracy and ability to account for the range of distinctions we need to make in this sphere but there is no basis for rejecting them altogether simply on the ground that they fail EP. While critical compatibilism certainly requires *some* (plausible) interpretation of conditions φ^* —and this remains a contested matter among them—the relevant standard for this assessment is not that it both respects and satisfies EP.³⁶

V. The Free Will Problem and Free Will Pessimism

We are now in a position to assess what significance the conclusion of FWPA (i.e., that we should reject EP and accept free will pessimism) has for the free will problem itself. There are two ways of interpreting the free will problem, one that is broad and general in character and another that is narrower and more specific. We may present the free will problem in the broad and general manner as being concerned with the question concerning the nature and conditions of freedom and moral responsibility. Plainly, however, viewing the problem in this more open-ended fashion does not serve to capture the more specific features of the free will problem and the standards of solution that have been set for it. As has been explained, throughout most of the modern period, the parties to this debate have enjoyed a shared understanding of the background assumptions and aspirations that frame this problem and what they would accept as an adequate solution for it. They are agreed, more specifically, that EP, as associated with the assumptions and (optimistic) aspirations of the morality system, is essential to the structure of this problem and that it serves as the relevant standard for any acceptable solution that may be proposed. In this way, the solution that all the parties to the free will problem have sought and struggled to find (p.117) is to provide an account of conditions of freedom and responsibility that satisfies EP.³⁷ Even skepticism, which denies that any such account can be provided, accepts that this is the relevant problem that stands in need of a solution and that EP must be *respected* even if it cannot be satisfied. For the skeptic this remains the relevant requirement for any adequate theory of freedom and moral responsibility and the problem of free will must be interpreted in these terms. To *dispense* with EP, therefore, is not to propose an alternative solution to the free will problem, so conceived, but rather to *reject the problem altogether*—along with its grounding and motivating assumptions and aspirations. These are, however, the very steps that FWPA takes and, therefore, free will pessimism should not be understood or interpreted as any kind of solution to the free will problem (i.e., given this stricter, narrower interpretation).

On the narrow understanding of the free will problem, all the parties involved in this dispute are agreed that free will pessimism cannot even be considered an available *candidate* for a solution, since it depends on rejecting the very assumptions that generate the problem in the first place—namely, EP. Considered from this perspective, the correct understanding of critical compatibilism is that it aims to *replace* the free will problem with an acceptance of free will pessimism, understood as a more truthful account of the human predicament. This predicament, along with its distinct pessimistic implications, is not a

problem to be solved but a predicament waiting to be recognized and acknowledged. Taking this step involves abandoning the evasions and fabrications of the morality system and the various modes of theorizing associated with it. When we abandon these assumptions and aspirations of the morality system, we do not solve the free will problem so much as cast it aside.

What, then, does free will pessimism contribute to the free will problem more broadly conceived? Even if we discard EP, and reject the free will problem as narrowly interpreted, we still cannot present free will pessimism as a *solution* to the free will problem more broadly conceived. Any solution along these lines would require a complete and convincing account of conditions φ^* . Although we may be fully persuaded by FWPA, we still face any number of significant issues and difficulties relating to the (contested) interpretation of conditions φ^* . While the narrow free will problem should be cast aside, and we can agree that accounts of conditions φ^* cannot and should not be rejected on the ground that they fail the standard of EP, critical compatibilism must still develop and defend its own preferred interpretation of conditions φ^* . Clearly, then, a defense of free will pessimism does not, by itself, provide us with a complete theory of critical compatibilism.

(p.118) Many suggested accounts of conditions φ^* may be found wanting and should be discarded. Even the most promising may still require further refinement or better articulation.³⁸

Having identified the limits of FWPA, we may now summarize its key claims and contributions. There are three interconnected components to this argument that are especially significant. First, the FWPA provides a diagnosis of why (orthodox) compatibilists have resisted the critical compatibilist approach and the free will pessimism associated with it. The fundamental source of resistance, it is argued, has its roots in the assumptions and aspirations of the morality system. Second, the FWPA makes clear that any adequate solution to the free will problem broadly conceived, as provided by some satisfactory account of conditions φ^* , will not deliver wholly optimistic conclusions about our human predicament in respect of these matters. On the contrary, free will pessimism is the troubling and difficult truth about our predicament and any credible account of conditions φ^* must recognize and acknowledge this. Third, and most importantly, the FWPA makes clear that any adequate and acceptable solution to the free will problem, broadly conceived, does not and cannot turn on a solution to the free will problem narrowly interpreted in terms of EP. The free will problem, more narrowly understood, has been generated by the faulty assumptions and illusory aspirations of the morality system. When

these assumptions and aspirations are set aside we are better placed to recognize and acknowledge the (difficult) truth about the human predicament, without the evasions and fabrications that are encouraged by the morality system.³⁹

References

Bibliography references:

Chisholm, Roderick. (1964). "Human Freedom and the Self". Reprinted in Gary Watson, ed. *Free Will*, 2nd edn. (Oxford: Oxford University Press (2003)), 24–35.

Dennett, Daniel. (1984). *Elbow Room: The Varieties of Free Will Worth Wanting*. (Oxford: Clarendon Press).

(p.119) Honderich, Ted. (2002). *How Free Are You?* 2nd ed. (Oxford: Oxford University Press).

Kane, Robert. (1996). *The Significance of Free Will*. Oxford: Oxford University Press.

Kane, Robert. (2007). "Libertarianism". In Fischer, Kane, et al., 5–43.

McKenna, Michael. (2012). *Conversation and Responsibility*. (New York: Oxford University Press).

Nagel, Thomas. (1976). "Moral Luck". Reprinted in Russell & Deery (2013), 31–42.

Pereboom, Derk. (2001). *Living Without Free Will*. (Cambridge: Cambridge University Press).

Pereboom, Derk. (2007). "Hard Incompatibilism". In Fischer, Kane, et al., 85–125.

Russell, Paul. (2000). "Compatibilist Fatalism". Reprinted in Russell & Deery (2013), 450–68.

Russell, Paul. (2002). "Pessimists, Pollyannas and the New Compatibilism", in Robert Kane, ed., *The Oxford Handbook of Free Will*. (New York: Oxford University Press), 229–56.

Russell, Paul. (2008). "Free Will, Art and Morality". *Journal of Ethics* 12: 307–25.

- Russell, Paul. (2013). "Responsibility, Naturalism and "the Morality System". In D. Shoemaker, ed. *Oxford Studies in Agency and Responsibility*. (Oxford: Oxford University Press), 184-204.
- Russell, Paul. (2017). *The Limits of Free Will: Selected Essays*. (New York: Oxford University Press).
- Russell, Paul, and Oisín Deery, eds. (2013). *The Philosophy of Free Will: Essential Readings from the Contemporary Debates*. (New York: Oxford University Press).
- Shoemaker, David. (2015). *Responsibility from the Margins*. (New York: Oxford University Press).
- Smilansky, Saul. (2000). *Free Will and Illusion*. (Oxford: Oxford University Press).
- Strawson, P. F. (1962). "Freedom and Resentment". Reprinted in Russell & Deery (2013), 63-83.
- Strawson, Galen. (1994). "The Impossibility of [Ultimate] Moral Responsibility". Reprinted in Russell & Deery (2013), 363-78.
- Taylor, Richard. (1959). "Determinism and the Theory of Agency". Reprinted in R. Shafer-Landau, ed., *Ethical Theory: An Anthology*. 2nd ed. (Oxford: Wiley-Blackwell, 2013), 308-11.
- Vargas, Manuel. (2007). "Revisionism". In Fischer, Kane, et al. (2007), 126-65.
- Vargas, Manuel. (2013). "How to Solve the Problem of Free Will". In Russell & Deery (2013), 400-16.
- Wallace, R. Jay. (1994). *Responsibility and the Moral Sentiments*. (Cambridge, MA: Harvard University Press).
- Waller, Bruce. (2011). *Against Moral Responsibility*. (Cambridge, MA: MIT Press).
- Watson, Gary. (1987). "Responsibility and the Limits of Evil." Reprinted in Russell & Deery (2013), 84-113.
- Watson, Gary. (1999). "Soft Libertarianism and Hard Compatibilism." Reprinted in G. Watson, *Agency and Answerability*. (New York: Oxford University Press (2004)), 197-215.

Williams, Bernard. (1976). "Moral Luck." Reprinted in B. Williams, *Moral Luck*. (Cambridge: Cambridge University Press (1981)), 20–39.

(p.120) Williams, Bernard. (1985a). *Ethics and the Limits of Philosophy*. With a commentary by A. W. Moore and a foreword by Jonathan Lear. (London and New York: Routledge (Routledge Classics edition 2011)).

Williams, Bernard. (1985b). "How Free does the Will Need to be?". Reprinted in B. Williams, *Making Sense of Humanity*. (Cambridge: Cambridge University Press (1995)), 3–21.

Williams, Bernard. (1993). *Shame and Necessity*. (Berkeley, CA: University of California Press).

Notes:

(¹) Williams (1985b: 19).

(²) Dennett (1984: ch. 1).

(³) Dennett (1984: 169).

(⁴) Kane (1996: 80).

(⁵) See, for example, Pereboom (2001, 2007), Honderich (2002), and also Waller (2011). This outlook may be compared, by analogy, to the Epicurean attitude to skepticism about immortality—i.e., mortality is not as depressing or disturbing as immortalists generally suppose.

(⁶) Some philosophers have argued that accepting the truth of skepticism about freedom and moral responsibility would be so damaging and depressing that we should encourage the illusion that we are free and responsible agents. See Smilansky (2000). Similarly, philosophical resistance to the skeptical conclusion is often motivated by the thought that this would be a "painful conclusion to accept" (Taylor 1959).

(⁷) Nagel (1976: 32–3).

(⁸) Dennett (1984: Ch. 4).

(⁹) Nagel (1976: 39).

(¹⁰) Nagel (1976: 40).

(¹¹) Dennett (1984: 104–5, 129).

(¹²) Russell (2002: 233–42), and also Russell (2000).

⁽¹³⁾ Whether we attach the label of “fate” to this concern is merely a verbal matter—the substantial concern or issue remains with us.

⁽¹⁴⁾ Dennett (1984: 94).

⁽¹⁵⁾ Nagel (1976: 38).

⁽¹⁶⁾ Williams (1985a: Ch. 10).

⁽¹⁷⁾ Williams (1985a: 217): “The purity of morality itself represents a value. ...”

⁽¹⁸⁾ Wallace (1994: 39–40, 64–6).

⁽¹⁹⁾ Wallace (1994: 39–40, 64–5), Galen Strawson (1994: 366).

⁽²⁰⁾ Dennett (1984): “... my conclusion is optimistic: free will is not an illusion, not even an irrepressible and life-enhancing illusion” (169). For criticism and doubts about Dennett’s optimism see Russell (2002: 249–51).

⁽²¹⁾ This is a prominent theme in many optimistic skeptical views about moral responsibility. See, for example, Pereboom (2001, 2007), and also Waller (2011).

⁽²²⁾ See, for example, Williams’ observation that “skepticism about the freedom of morality from luck cannot leave the concept of morality where it was ...” (1976: 39). See also Williams (1985a: 53, 170, 1985b: 19–20, and esp. 1993: 126, 158–67).

⁽²³⁾ Williams (1985a: 43, 73, 215–17). Related to this, Williams also emphasizes the way in which the morality system places particular emphasis on the importance of voluntariness and blame in this context.

⁽²⁴⁾ In relation to this see Russell (2008). See also P. F. Strawson’s remarks noting that inside the web of reactive attitudes and feelings “there is endless room for modification, redirection, criticism, and justification” (1962: 81).

⁽²⁵⁾ One such troubling case that helps to illuminate free will pessimism, and highlights the way in which it contrasts with theories constructed around the demands of EP, is that of “Robert Harris” (who is discussed in some detail in Watson (1987)).

⁽²⁶⁾ My discussion of libertarianism in this section (relating to premises #8 and #9) has been compressed but a more detailed account is presented in an extended version of this paper in Russell (2017).

(²⁷) Taylor (1959: 310).

(²⁸) Chisholm (1964: 34).

(²⁹) Kane (2007: 23, 27). See also Strawson (1962: 83), who famously describes libertarianism of this general kind as relying on “obscure and panicky metaphysics.”

(³⁰) Kane (1996). The soft/hard distinction in relation to both libertarianism and compatibilism is presented in Watson (1999).

(³¹) See, for example, Pereboom (2001: 50–4) and Pereboom (2007: 103–10) for a clear statement of this general line of criticism to event-causal theories.

(³²) Wallace (1994: 54–6).

(³³) Wallace (1994: 39–40, 64–6). For criticism of Wallace’s “narrow” interpretation of responsibility as framed in terms of the apparatus of “the morality system” see Russell (2013).

(³⁴) See, for example, Strawson (1994), which was originally published as “The Impossibility of Moral Responsibility” but subsequently published under the more restricted title “The Impossibility of *Ultimate* Moral Responsibility” (my emphasis). The insertion of “ultimate” marks a drift from global to local skepticism and, therefore, a considerable shift in the significance and scope of the conclusion that is being drawn from his “basic [skeptical] argument”. A similar slide or ambiguity can, perhaps, also be detected in Pereboom’s contributions, where he discusses “analogues” of the reactive attitudes that survive the skeptical critique that he advances (2001: 97, 199–200, 2007: 118–20). (See also n. 35.)

(³⁵) It may be argued that Pereboom’s “hard incompatibilism” is liable to this line of criticism since, on the face of it, hard incompatibilism may be read as a combination of *global* skepticism and optimism. There is, however, an alternative, weaker reading that presents hard incompatibilism as only *locally* skeptical about “basic desert” views of freedom and responsibility (i.e., of the kind encouraged by the morality system), while still rejecting global skepticism of any kind. On the weaker reading both hard compatibilism and critical compatibilism may well converge and be able to accommodate each other in these important respects (i.e., since they are both locally skeptical but also reject global skepticism). Nevertheless, from the perspective of critical compatibilism, the crucial point is that rejecting the morality system and its particular interpretation of freedom and moral responsibility

does not leave us in a wholly comfortable situation that licenses complacency or optimism. On the contrary, when we abandon EP, we must accept free will pessimism and acknowledge the troubling predicament associated with it.

⁽³⁶⁾ Critical compatibilism, so described, might be viewed as simply a version of “revisionist” theory of freedom and responsibility, for example Vargas (2007, 2013). There are, however, a number of reasons for resisting this suggestion, including the fact that some *orthodox* compatibilist theories also present themselves as “revisionary” (see, for example, Dennett 1984: 19). In general, what matters for a proper assessment of any proposed revisionary theory, as it concerns critical compatibilism, is, first, where it stands in relation to EP and, second, what stance it takes with respect to the metaphysical attitudes involved. A “revisionary” theory may or may not reject EP and, even if it does, it may or may not endorse the particular pessimistic attitudes that are drawn from this by way of free will pessimism (as some theorists may regard dispensing with EP as entirely untroubling, if not liberating). Suffice it to note that most contemporary revisionist projects follow in the neo-Skinnerian tracks of Dennett’s optimistic pragmatism.

⁽³⁷⁾ As we have noted, this includes (orthodox) compatibilists, such as Wallace (1994: 64–6).

⁽³⁸⁾ For two important and impressive recent contributions that offer (compatibilist-friendly) accounts of moral responsibility see McKenna (2012) and Shoemaker (2015).

⁽³⁹⁾ This paper has gone through multiple drafts and has been presented to a number of different audiences over a period of several years. This includes presentations at Leiden (2009), Budapest/CEU (2009), Oxford (2010), Arizona (2012), Delaware (2013), Montreal (2014), Chicago (2014), Aarhus/Danish Philosophical Association (2015), Gothenburg (2015), and NOWAR (2015). I am grateful to many individuals who were present on those occasions for their valuable comments and criticisms. Although I am not able to thank them all, I am especially grateful to John Fischer, Michael McKenna, Derk Pereboom, Jesse Prinz, Saul Smilansky, Galen Strawson, and Daniel Telech.



Access brought to you by: **Lund University Libraries**