

9

Free Will and Affirmation: Assessing Honderich's Third Way

Paul Russell

In the third and final part of his *A Theory of Determinism* (1988), Ted Honderich addresses the fundamental question concerning “the consequences of determinism” (1988, II, 4).¹ That is, what follows if determinism is true? This question is, of course, intimately bound up with the problem of free will and, in particular, with the question of whether or not the truth of determinism is compatible or incompatible with the sort of freedom that is supposed to be required for moral responsibility. It is Honderich’s aim to provide a solution to “the problem of the consequences of determinism,” and a key element of this is his attempt to collapse the familiar Compatibilist/Incompatibilist dichotomy. More specifically, Honderich offers us a third way—the response of “Affirmation” (2002, 125–126). Although his account of Affirmation has application and relevance to issues and features beyond

P. Russell (✉)

[AUI](#) University of British Columbia, Vancouver, BC, Canada

University of Gothenburg, Gothenburg, Sweden

© The Author(s) 2017

G. Caruso (ed.), *Ted Honderich on Consciousness, Determinism, and Humanity*,
Philosophers in Depth, DOI 10.1007/978-3-319-66754-6_9

18 freedom and responsibility, my primary concern in this chapter will be
19 to examine Honderich's theory of "Affirmation" as it concerns the free
20 will problem as it is generally understood.

21 9.1

AU2

22 The first part of Honderich's overall position in *A Theory of Determinism*
23 is concerned with "the question of whether there does exist a conceptually
24 satisfactory determinist theory of our lives" (1988, 4). The second
25 part addresses the question of whether the theory as articulated in the
26 first part and argues that it is well supported. The core of this theory, as
27 Honderich interprets it, is expressed in these terms:

28 **determinism** is only a view of our own nature—in essence, the view that
29 ordinary causation is true of us and our own lives, that in our choosing
30 and deciding we are subject to causal laws. In this use of the word, deter-
31 minism comes to no more **that** a yes answer to the question of whether
32 we are in one fundamental way like plants or machines. Determinism in
33 this sense does not include or imply an answer to the question of whether
34 we are free or not. That question, maybe surprisingly, is left pretty well
35 untouched. (2002, 3)



AU3

36 The account of the theory of determinism is developed by Honderich
37 largely in terms of a theory of causation and laws, along with an accom-
38 panying theory of mind and action. The theory of causation is one that
39 holds that causes necessitate their effects, such that, given the occur-
40 rence of the former, whatever else happens, the latter will also occur
41 (2002, 14–15). Causal relations and natural laws are understood as in
42 terms of "regularities or as nomic connections" (2002, 15–16). Given
43 this general account of causation, the question arises of whether or not
44 "our lives are a matter of effects ... events that really have to happen
45 because of earlier causal circumstances" (2002, 21). To answer this
46 question, as it concerns our choices and actions, Honderich advances
47 his theory of mind and action.

9 Free Will and Affirmation: Assessing Honderich's Third Way

There are three basic elements to Honderich's account of mind and action. The first is what he calls Mind-Brain Determinism. This concerns the suggestion that each mental or conscious event, including choosing and deciding, is nomically related to neural events in such a way that the "neural event by itself or together with some other non-mental thing necessitated the mental one" (2002, 63). The second component is "Intuition Determinism," which maintains that nomically related neural and mental events have their causal origins in bodily (and environmental) events that involve no mental events. The third, and last, component in this general account of mind and action is "Action Determinism," which claims that each and every action is the effect of an "active intention," what is sometimes referred to as a volition or act of willing (2002, 59–60). The details of these three elements of Honderich's theory of mind and action involve, of course, a number of controversial issues but for our present purposes what matters is that they serve as the background set of proposals and assumptions that serve to articulate the general theory of determinism and bring us directly to the problem of the "consequences" (so interpreted). It is worth noting that Honderich does not assert simply "that determinism is true." What he claims is that it is "strongly supported, and that certainly it *has not been shown to be false*" (2002, 90 [original emphasis]). Honderich also allows that there may be some "micro-indeterminism" at the level of small particles it is "not amplified into indeterminism at the macro-level" (2002, 74–76). Even if micro-determinism is true, we are still left with strong support for the truth of a "macro-determinism" or "near-determinism" (2002, 90).

Throughout his account of a deterministic theory of mind and action, Honderich makes clear that there is an opposing view—namely, what he calls the "philosophy of Free Will," which presupposes the truth of indeterminism (2002, 2, 4, 35, 41–42, 49, 69, 76). What philosophers of Free Will require is that responsible agents are originators of their actions. If we are genuinely free, each of us must have "a kind of personal power to *originate* choices and decisions and their actions" (2002, 2). It will not suffice for us to be mere "initiators" of our actions in a manner that is consistent with these events to be (necessitated) effects via a causal chain or process (2002, 3). The guiding aim behind the idea of Free Will,

83 Honderich argues, is that we need to understand responsibility *in a cer-*
84 *tain way*. What this requires, above all, is that we are “able now to choose
85 differently from how we do, given the present and ourselves exactly as
86 they are and the past exactly as it was” (2002, 41–42; and cp. 98, 109,
87 117). This form of absolute responsibility not only requires the falsity of
88 determinism it also requires some sort of “ongoing entity” that possesses
89 the “active power” required to produce or not produce a given action
90 (2002, 42, 49). If we accept the truth of (near) determinism, then we
91 must reject the idea of Free Will, so understood.

92 What is the significance of having to abandon or discard the idea of
93 Free Will? Honderich analyses this issue within wider fabric of what he
94 refers to as a set of “life-hopes.” Life-hopes involve our attitudes to a range
95 of features of the world that we care about, particularly as this concerns
96 how our future activities may affect or influence how our lives unfold. In
97 respect of these matters, we may be fortunate or unfortunate and, depend-
98 ing on this, we will *feel* a general attitude of a positive or negative kind.
99 An attitude is an “evaluative thought” about something that we either
100 approve or disapprove of. The question that presents itself in relation to
101 determinism is, therefore, whether our life-hopes are in some relevant way
102 threatened or damaged if the thesis of determinism is true.

103 It is with respect to this matter that the loss of Free Will is of some
104 concern. The idea of Free Will, as we noted, involves a conception of
105 ourselves as agents that are true originators of our actions. There is,
106 Honderich maintains, a set of life-hopes that rests on this stronger con-
107 ception of free agency (2002, 93, 111–112). We want to be able not just
108 to achieve success but to *earn* it and we want our achievements, whatever
109 they are, to be more than just a causal product of our inherent nature and
110 environment. While this is of importance to us if determinism is true, we
111 do not have it (2002, 104). This situation will license an attitude of dis-
112 may about our predicament and circumstances so understood. It is for
113 this reason, Honderich suggests, “that many people have found deter-
114 minism to be a black thing” (2002, 95; and cp. 104).

115 Life-hopes are not, however, the only things that rest on the concept of
116 freedom understood in terms of origination. Among other things that
117 depend on this notion of initiated action as originated are certain attitudes
118 or personal feelings to other people on the basis of their conduct toward
119 us (e.g. resentment and gratitude). These personal feelings and reactions

9 Free Will and Affirmation: Assessing Honderich's Third Way

are closely bound up with our *moral* feelings and responses—an issue that has been the primary focus of attention in the free will debate. What is crucial here is that the concept of freedom as originated action is essential to moral responsibility. The moral standing of an individual, on this account, rests on the assumption that they have not just acted voluntarily but possess the power to originate their actions and could have acted otherwise in the same circumstances. If these metaphysical foundations of agency are threatened, the whole fabric of morality will collapse, including our understanding of retributive justice and (deserved) punishment (2002, 101). Our sense of dismay, in face of the apparent truth of determinism, is felt most strongly in relation to this issue.

It is at this juncture that the split between Compatibilism and Incompatibilism becomes relevant. The analysis provided so far is one that turns on a concept of freedom (i.e. origination) that the incompatibilist embraces and endorses as the only true or genuine account of freedom that is relevant to these modes of concern. From the perspective of compatibilism, however, this entire analysis is mistaken. It is mistaken because what we really mean by freedom is a matter of voluntariness—not origination (2002, 96, 98). Free actions have a certain kind of causal history and are, as such, effects of a particular kind (2002, 96). When we secure a proper grasp of the concept of freedom, in terms of voluntariness, we can draw all the relevant distinctions we need concerning moral responsibility (2002, 98–99). There is, therefore, according to the compatibilist no conflict between freedom and responsibility and the truth of determinism. Even if determinism is true, nothing changes and nothing we have reason to care about is threatened. Instead of dismay, the attitude of the compatibilist is one of “satisfied intransigence” (2002, 97).

It is Honderich's central claim that neither of these approaches offers a satisfactory response to the likely truth of determinism. The oppositions we encounter between these two sides leave us in an unsatisfactory and unstable situation that demands a new response to determinism.

We need to get into a different way of feeling about determinism. We need to come to a response that takes into account not only its truth, and the two sets of attitudes, but also the two responses we have in the first instance, dismay and intransigence. So the final upshot, if we are to be successful, will partly be a response to the two initial responses. (2002, 122–123)

156 Honderich's response to compatibilism and incompatibilism is that they
157 are *both mistaken*. Both, moreover, are mistaken over several overlapping
158 claims that they are agreed about. The first of these is that we possess
159 some single, ordinary idea of freedom and that the alternative conception
160 is either inadequate or incoherent (or both). The truth about our situa-
161 tion, Honderich maintains, is that we plainly have two distinct ideas of
162 free choice and action and that both compatibilists and incompatibilists
163 are wrong to dismiss the alternative view as irrelevant to our appreciation
164 of the consequences of determinism. Any proposed solution along these
165 lines will inevitably be one-sided and incomplete and will fail to identify
166 properly the (inescapable) split we experience in our dual responses and
167 attitudes (i.e. *both* dismay *and* intransigence).

168 Honderich argues that both parties in this debate mistakenly suppose
169 "that the problem of the consequences can be settled by logical, intellec-
170 tual, philosophical or linguistic means as traditionally conceived" (2002,
171 114) This approach, he suggests, over-intellectualizes the whole problem.
172 The real problem is that we have two sets of attitudes, rooted in two sets
173 of desires, and they issue in the divergent responses we have considered—
174 dismay and intransigence. Neither response, taken by itself, is entirely
175 satisfactory. Compatibilist intransigence comes across as mere "bluff" and
176 requires us to suppress and ignore the other side of the equation. On the
177 other hand, simple dismay also ignores essential and inescapable features
178 of ordinary life and our social experience. The relevant problem is how
179 this "Attitudinarian instability and discomfort can be overcome and the
180 two parties reconciled" (2002, 125–126).

181 Honderich's solution to this problem takes the form of what he calls
182 "Affirmation" (2002, 126). What this requires is that, assuming the truth
183 of determinism, we must

184 try to give up whatever depends on thoughts inconsistent with it. Above all
185 we have to try to accept the defeat of certain desires. This is bound up with
186 *trying to be happier about, or more reconciled to, the desires in which we can*
187 *persist, the ones consistent with determinism.* (2002, 125)

188 This means in the first place, in accepting that "our attitudes involving
189 voluntariness cannot really allow us to be intransigent." If determinism is

9 Free Will and Affirmation: Assessing Honderich's Third Way

true, we cannot claim that “nothing changes” and that this “leaves things just as they are” (2002, 97, 99, 125). At the same time, making these concessions “need not give rise to dismay, taking everything as wrecked” (2002, 125). By this route we achieve some balance between dismay and intransigence, whereby we recognize and accept that some things important to us are discredited or lost, to the extent that they are based on assumptions about origination, but that there is also a great deal that we care about and that which matters to us remains untouched and survives.

Honderich summarizes his account of Affirmation as follows:

To put the proposal in a nutshell, our new response should be this: trying *by various strategies to accommodate ourselves to the situation we find ourselves in—accommodate ourselves to just what we can really possess if determinism is true, accommodate ourselves to the part of our lives that does not rest on the illusion of Free Will.* We can reflect on what is perhaps the limited worth of what we have to give up, consider the possible compensations of a belief in determinism, take care not to underestimate what we can have, and consider a certain prospect having to do with genuine and settled belief in determinism. (2002, 126 [emphasis in original])

To embrace this response of “Affirmation” is, Honderich suggests, to adopt “a philosophy of life,” one which consists in feelings that provide us with some support and as much satisfaction as the truth will allow. Although Affirmation rejects all forms or reliance on illusions of any kind, it nevertheless rejects any (pessimistic) suggestion that determinism leaves us “defeated” or without any substantial and significant sources of consolation. This is not just a matter of “putting a good face on things” but of recognizing gains as well as losses that come with the belief in determinism (2002, 131). Among the gains is that we are in a position to withdraw from retributive sentiments and practices that are grounded in our illusory belief in Free Will. This is all achievable without the world going “cold,” or leaving no scope for personal emotions and feelings, or losing all sense of achievement and meaning (see Pereboom 2001, 2013, 2014). The upshot of all this is that the response of Affirmation is one that rejects undiluted “dismay” or “intransigence” but offers, instead, a *blend* or *mix* that avoids the one-sided, monochromatic alternatives that have generally been advanced.

9.2 II

226 Having reviewed the essential features of Honderich's theory of
 227 "Affirmation," we can now ask if this is indeed, as Honderich claims, the
 228 solution we have been seeking for "the problem of the consequences of
 229 determinism" (2002, 133). The short answer to this—as with the theory
 230 itself—is in some ways "Yes" and, in other ways, "No." Let us begin with
 231 what seems to be right and illuminating about "Affirmation." The best
 232 way to appreciate the significance of Affirmation is by locating it within
 233 the matrix of other available positions on offer. Affirmation is one of several
 234 positions on this subject that involve significant modifications and
 235 amendments to the more familiar classical accounts. (As Honderich
 236 points out, neither the Compatibilist or Incompatibilist tradition is
 237 "absolutely uniform" (2002, 110).) Although several of these theories
 238 contain overlapping or common elements, each takes a different turn on
 239 some key issues—and Affirmation needs to be distinguished from them
 240 in respect of these elements.

241 Affirmation is most obviously opposed to the idea of Free Will,
 242 understood in terms of the form of origination that libertarians seek
 243 to secure. In this respect, Determinism and Affirmation plainly imply
 244 skepticism about the metaphysics of libertarian agency—which is,
 245 according to Affirmation, a persisting source of dismay. However,
 246 although Affirmation is skeptical about Free Will and origination, it
 247 rejects *complete*, global skepticism about freedom and responsibility.
 248 The reason for this is that it rejects the suggestion that origination is
 249 our (sole) true idea of freedom. According to Affirmation, voluntariness
 250 serves as one kind of freedom and provides a basis for surviving
 251 credible forms of responsibility based upon it. It follows from this
 252 that any *unqualified* form of dismay or pessimism cannot be supported
 253 or sustained, simply because origination is incompatible with the truth
 254 of determinism. Affirmation, therefore, rejects "incompatibilist"
 255 arguments about freedom and responsibility, and any unqualified
 256 pessimism that it may be taken to license on the ground that they
 257 fail to accommodate the dual nature of our concepts of freedom and
 258 responsibility (see, e.g. Strawson 1994).

9 Free Will and Affirmation: Assessing Honderich's Third Way

Not all skeptics about freedom and moral responsibility take themselves to be committed to dismay or pessimism as a metaphysical attitude that flows from the truth of determinism. An important recent development in the free will debate has been an effort, by Derk Pereboom, Gregg Caruso and others, to defend “Hard Incompatibilism” or “Optimistic Skepticism,” which holds that the only form of freedom that grounds desert-based theories of responsibility is origination and that, although this is impossible, no deep or unqualified pessimism flows from this. On the contrary, there are, they claim, significant *benefits* to this skeptical outlook (see, e.g. Pereboom 2001; Waller 2011; Caruso, forthcoming). There are certainly some important affinities between Honderich's Affirmation and “Hard Incompatibilism,” but, as Honderich shows, there remain important points of divergence (2002, 143). One aspect of this is that Affirmation remains firmly committed to the stance of dismay (or pessimism)—rooted in skepticism about the idea of Free Will—which is a stance that Hard Incompatibilism aims to discredit or minimize. The crucial point here is that Affirmation refuses to deny the source of dismay that determinism generates, as grounded in skepticism about the (Incompatibilist) idea of Free Will or origination.

One further theory that we should consider in this context is “Illusion,” as advanced and defended by Saul Smilansky (2000, 36–38). As with Hard Incompatibilism, Smilansky's “Illusionism” bears some family resemblance to Affirmation. One important feature they share is a “dualist” view about the nature of freedom and responsibility, which denies that we have just one conception which renders either Compatibilism or Incompatibilism true or false. Where they diverge, however, is that “Illusionism” takes the importance of origination to be so deep and pervasive in our ethical lives that for practical reasons we should encourage and promote *belief* in origination—even if we have some theoretical reason for doubting it (e.g. evidence of the truth of determinism). For Illusionism, the truth of determinism and abandoning the idea of free will would be so catastrophic and disastrous for our ethical and social lives that we must reject any option or theory that would encourage skepticism about the idea of Free Will. This is a view that Honderich, explicitly rejects. The philosophy of life that Affirmation embraces insists that the

294 path of Illusion is neither necessary nor desirable and that Affirmation
295 secures all that is needed to sustain and support a worthwhile life and a
296 viable ethical community (2002, 126, 131–132).

297 Where does this taxonomy of Affirmation in relation to other alterna-
298 tive views leave us? In my view, Affirmation not only presents a distinc-
299 tive stance and position on this issue, on all the points and issues that
300 separate it from its alternatives mentioned above, it generally takes the
301 right view. Affirmation is correct, for example, to reject “Monism” about
302 the concept of freedom as it concerns moral responsibility. Related to
303 this, it is right to reject any unqualified skepticism about freedom and
304 responsibility, along with any unqualified pessimism or dismay that may
305 be based on it. Finally, I would also agree that Affirmation is right to
306 repudiate “Illusionism” as a way of dealing with the “probable” truth of
307 determinism, insist instead on being truthful about our human predicam-
308 ent. In several respects Affirmation shares some of the key merits and
309 strengths in P.F. Strawson’s influential contribution to this topic in his
310 paper “Freedom and Resentment.”² Along with Strawson, Affirmation
311 places emphasis on moral sentiments or reactive attitudes in accounting
312 for moral responsibility. Both aim to “reconcile” the main parties in this
313 dispute by means of accepting some of their claims and not others and
314 present a position that gives us a recognizable picture of human agents as
315 part of the natural world. However, what does separate Affirmation from
316 Strawson’s theory is that it does not claim to “leave things just as they are”
317 (2002, 99). Affirmation insists on recognizing the real losses, as well as
318 the gains, that come with this philosophy of life (2002, 126, 131). In all
319 these respects, so it seems to me, Affirmation is firmly on the right track
320 and highly illuminating.

321 Having indicated what makes Affirmation distinctive, illuminating,
322 and credible, let me now consider some other important matters in
323 respect of which it is less convincing. In order to do this, I want to com-
324 pare Affirmation with another perspective on the free will problem—a
325 view I will call Free Will Pessimism (Russell 2017a). The label Free Will
326 Pessimism does not involve a metaphysical attitude that is based on
327 some form of free will skepticism. On the contrary, whereas free will
328 skepticism is the view that our vulnerability to conditions of fate and
329 luck serve to discredit our view of ourselves as full and responsible agents

AU4



9 Free Will and Affirmation: Assessing Honderich's Third Way

(e.g. as implied by the truth of determinism), Free Will Pessimism *rejects* free will skepticism. The basis of its pessimism rests with the assumption that we *are* free and responsible agents who are, nevertheless, subject to fate and luck. According to Free Will Pessimism, all the major parties and positions in free will debate (both Compatibilists and Incompatibilists) are committed to modes of evasion and distortion regarding our human predicament in respect of agency and moral life. The question that arises, therefore, is whether or not the “dualism” involved in Affirmation, in respect of both its understanding of two concepts of freedom and the associated split in our metaphysical attitudes, effectively identifies and overcomes these modes of evasion?

In order to describe the alternative picture that Free Will Pessimism presents of the free will problem as it relates to the consequences of determinism, we need to consider, first, 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).
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 interpretation of conditions φ can satisfy EP and avoid the skeptical conclusion, although this requires the

AU5

AU6



364 falsity of determinism. Skeptics maintain that there is no avoidable set of
365 conditions φ that serve to satisfy EP and, hence, the skeptical conclusion
366 goes through either way—whether determinism is true or false.

367 Proponents of BEA are entirely justified in claiming that compatibil-
368 ists have consistently adhered to EP and aimed to satisfy it. What compatibilists
369 have denied is premise #4, the claim that compatibilism fails to
370 satisfy the standard set by EP (premise #3). Let us consider, then, the
371 classical compatibilist argument that is launched against premise #4, an
372 argument aiming to show that agents who satisfy suitably interpreted
373 compatibilist conditions (φ^*) are not subject to fate and luck (i.e. condi-
374 tions β). The core feature of this argument is that the incompatibilist
375 claim (premise #4) relies on a basic confusion between fatalism and
376 determinism. More specifically, if we properly interpret conditions β (i.e.
377 as conditions β^*), then premise #4 is groundless. Fatalism is the doctrine
378 that all our deliberations and actions are *causally ineffective* and make no
379 difference but nothing about the thesis of determinism implies that this
380 is the universal condition (see, e.g. Ayer 1954, 22–23; and especially
381 Dennett 1984, 104–105, 129).

382 This line of response, aimed at refuting premise #4, may be found
383 doubly unconvincing. First, appealing to this distinction between deter-
384 minism and fatalism is a shallow and evasive understanding of incom-
385 patibilist concerns. The relevant issue is not about the causal influence *of*
386 the agent but rather the causal influence *on* the agent. On the assumption
387 of determinism, however complex the mechanisms or capacities involved,
388 the ultimate source or origin of conduct and character is external to the
389 agent and not within the agent's control or influence. For this reason, we
390 need to distinguish “contributory fatalism,” which concerns the universal
391 causal impotence of agents, from “origination fatalism,” which concerns
392 the causal source or origins of an agent's conduct and character and the
393 limits of control over this. Whereas determinism does not imply univer-
394 sal contributory fatalism, it does imply universal origination fatalism—
395 and it is this that is found particularly troubling (Russell 2000, 2017a).³

396 The second objection flows from the first. Given the issue of origina-
397 tion fatalism and the limits of control over conduct and character, as
398 generated on compatibilist models, we run directly into worries about
399 moral luck. The general worry here is about agents being subject to moral

evaluation (i.e. reactive attitudes etc.) in ways that are sensitive to factors they do not control (Nagel 1976). Granted that it is intuitively unjust to hold agents responsible for aspects of their conduct and character that they do not control (as per EP, premise #3), conditions of freedom and responsibility cannot be sustained in these circumstances where an agent is subject to fate and luck along the lines described. The familiar compatibilist line of response to this, consistent with much contemporary compatibilist thinking, is that human agents are “not just lucky,” we are “skilled self-controllers” (Dennett 1984, 94). Once again, however, 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 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 no agent ultimately has control over.

On this account, the free will problem, as generally understood, turns on the assumption that EP is correct and that the most effective compatibilist strategy depends on refuting premise #4. Incompatibilists claim that this cannot be done and that determinism implies skepticism about freedom and responsibility. There is, however, an alternative strategy we may call “Critical Compatibilism.” Critical Compatibilism *accepts* premise #4 and agrees with incompatibilists that orthodox compatibilist attempts to refute are shallow and evasive and, as such, fail. At the same time, critical compatibilists reject the skeptical conclusion because they reject EP/premise #3. According to Critical Compatibilism, *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 taking this view, critical compatibilism maintains that any plausible form of compatibilism must accept Free Will Pessimism, which allows that free and responsible agents may still be subject to significant forms of fate and luck (contrary to the requirements of EP).

An important feature of Critical Compatibilism is the particular set of metaphysical attitudes that this stance naturally licenses or occasions. In circumstances where EP is not satisfied, we have (deep) reasons for being

436 “troubled” or “disconcerted” by our predicament as this relates to human
437 ethical life and moral agency. Even if we are “fortunate” in **this** particular
438 ethical trajectory our lives may take, there is no basis—as incompatibilists
439 rightly insist—for an easy optimism when fate and luck intrude into our
440 ethical lives and the way we may exercise our moral agency. The crucial
441 point, in relation to Critical Compatibilism, is that a pessimism (or sense
442 of “dismay”) of this nature is not rooted or grounded in skepticism about
443 free will and moral responsibility. On the contrary, it presupposes that we
444 *reject* any skepticism of this kind, since the form of pessimism that is
445 occasioned depends on viewing ourselves and others as agents who are
446 free and responsible but, nevertheless, subject to fate and luck in the exer-
447 cise and operation of our moral capacities.

448 With this understanding of the central features of Free Will Pessimism
449 in place, we may now turn back to Affirmation and ask where it stands in
450 relation to Free Will Pessimism? We can begin by asking, more specifi-
451 cally, where Affirmation stands on the issue of whether to accept or reject
452 EP (premise #3 of BEA)? Clearly, Affirmation may go in one or other of
453 two directions. Affirmation may be interpreted—and, on the face of it, is
454 most plausibly interpreted—as offering a “dualist” response to EP which
455 holds that it is *both* satisfied and not satisfied, depending from which side
456 of the Compatibilist/Incompatibilist fence we are viewing it. Given the
457 truth of Determinism and the consequent impossibility of origination or
458 Free Will, Affirmation will agree with Incompatibilists that EP cannot be
459 satisfied in *terms of this concept*—and this will license our sense of dismay
460 based on skepticism about freedom and moral responsibility (as sug-
461 gested by BEA). This is, however, only one side of the mixed dualism of
462 Affirmation. Since we have another concept of freedom, understood in
463 terms of voluntarism, which allows for a different set of implications rela-
464 tive to freedom and responsibility, EP may be satisfied by the relevant set
465 of Compatibilist standards. This is the case, if we accept the orthodox
466 compatibilist project, which turns on a refutation of premise #4 (and an
467 acceptance of EP).


468 If Affirmation is interpreted in the manner described above, then
469 neither of its dualist components involves rejecting or discrediting EP
470 and embracing any form of Free Will Pessimism. The Incompatibilist
471 side holds that EP cannot be satisfied but must be respected (hence the
472 skeptical conclusion follows). The Compatibilist side holds that EP can



be satisfied, as long as we do not conflate modes of freedom and responsibility secured by voluntariness with those that are based on origination or Free Will. According to Affirmation, both these views can be held together, without contradiction. The two views, along with their accompanying set of attitudes, can be simultaneously held and are insulated from each other because both rest on ideas and concepts that are equally firmly rooted in ordinary life. For this reason, given the truth of Determinism, our reflections about EP will generate a measure of both dismay and intransigence—such that the truth of Determinism neither leaves our lives in ruins nor leaves everything unchanged and where it was. The crucial point remains, however, that although dismay attaches to the loss of origination, it does not generate any form of Free Will Pessimism, since, on both sides, Affirmation remains committed to *respecting* EP (if not satisfying it). Affirmation, so interpreted, denies the very possibility of Free Will Pessimism (i.e. it accepts EP as it appears in BEA). Another way of putting this point is that, along with the orthodox Compatibilist, Affirmation holds that Compatibilists should hold onto EP **but** reject premise number #4, which claims that Compatibilist conditions may be satisfied while agents are still subject to fate and luck. The Compatibilist element of Affirmation is, on this account, orthodox and not Critical Compatibilist in content. Like other views in the free will debate, Affirmation does not take the step taken by Critical Compatibilists and Free Will Pessimism, which involves discarding the entire ambition of satisfying EP or to find a “solution” conceived in these terms.

Although Affirmation may be interpreted in these terms, it is unsettled and open-ended enough, in relation to these matters, that another interpretation is available to it. Affirmation may accept, with Critical Compatibilism, that any plausible and credible form of Compatibilism must accept and embrace Free Will Pessimism. That is to say, the Compatibilist component of Affirmation would not only reject simple intransigence because Determinism implies skepticism about origination and the idea of Free Will, but also because Compatibilism, understood in terms of freedom and moral responsibility based on voluntarist views, has its own independent sources of pessimism. This distinct and independent source of pessimism is the claim that free and responsible agents (in Compatibilist terms) are still subject to significant modes of fate and

510 luck. All orthodox Compatibilist efforts to dismiss this (via refutations of
511 premise #4) are rejected as themselves evasive and shallow responses to a
512 more truthful account of the human predicament. On this interpretation,
513 therefore, there is nothing about Affirmation that *precludes* it from endors-
514 ing Free Will Pessimism. The difficulty with Affirmation, as presented, is
515 that despite its other insights, it fails to provide a clear assessment of these
516 issues relating to EP and Free Will Pessimism, leaving its Compatibilist
517 component *indeterminate* with respect to these important matters.

518 One reason why these matters are of such importance is that the sort
519 of “solution” we are looking for here depends on how we understand the
520 problem of Determinism in relation to the free will problem. Honderich
521 presents Affirmation as a solution to the problem of Determinism based
522 on its dual “Attitudinarian” components (2002, 104, 120–121, 122, 126,
523 133). He presents Affirmation as restoring a kind of balance and stability
524 in our divergent and contrary responses to the truth of Determinism and
525 as providing us with a “philosophy of life” that allows us to recognize and
526 accept certain losses while retaining sufficient resources to carry on with
527 some comfort and confidence. The difficulty we are faced with, however,
528 is to identify carefully and precisely the relevant persisting sources of dis-
529 may or pessimism and how exactly they relate to the Compatibilist side
530 of Affirmation. As presented, Honderich’s account seems to locate all rel-
531 evant sources of dismay with the loss of origination. If the Critical
532 Compatibilist is correct, this is a mistaken and inadequate analysis of the
533 problem. It is not just that we cannot satisfy the aim and aspirations of
534 EP in Incompatibilist terms but that we need to discard EP **altogether** as
535 a standard for an acceptable solution to the freewill problem (a matter 
536 which even the skeptic is mistaken about). The *cost* of rejecting or dis-
537 carding EP as a basic assumption for assessing proposed “solutions” to the
538 free will problem is that we must allow that free, responsible agents are
539 still subject to fate and luck—and this is something that is independently
540 disturbing and troubling and plainly cannot be based on any general
541 skepticism about freedom and responsibility. If Affirmation embraces
542 Free Will Pessimism, then it must abandon any suggestion that it is a
543 “solution” to the free will problem conceived in terms of the ambition to
544 *satisfy* EP on some relevant interpretation. The Free Will Pessimist claims
545 that what we have here is not a (puzzling) problem to be solved, but a

troubling predicament to be acknowledged. All the major parties in the free will debate—including orthodox Compatibilists—are resistant to the suggestion that free, responsible agents may still be subject to significant modes of fate and luck. What still needs to be clarified, however, is where Affirmation stands on this matter.

The foregoing problem is indicative of weaknesses in Affirmation's commitment to the two concepts/two attitudes formula at the heart of its analysis. The split between Compatibilism and Incompatibilism that serves as the framework of its analysis conceals what all the traditional parties share—namely, a commitment to standards of freedom and responsibility that satisfy EP. As we have noted, even the skeptic shares this standard, taking the view that EP must be respected, even if it cannot be satisfied. It is this *more fundamental conception* that is the real obstacle, not just for understanding the consequences of the truth of Determinism, but for understanding the truth about the human predicament with respect to agency, whether (near) Determinism is true or not. The question we need to ask, at this juncture, is why have *compatibilists* been so reluctant to embrace critical compatibilism and Free Will Pessimism? I have provided a more extended answer to this question elsewhere, but for now a brief summary will suffice (Russell 2017a). 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 relationship between the exclusion premise and “the morality system” and the second, related to the first, concerns the question of optimism.

With respect to the first point, the exclusion premise may be understood as an essential feature of what Bernard Williams calls “the morality system” (1985, Ch. 10). Among the various distinguishing features of “the morality system” is its insistence that moral responsibility, rightly understood, must somehow be capable of “transcending luck,” providing a purity that only genuine “rational” agency of some kind makes possible (Williams 1985, 217; see also Nagel 1976). Although orthodox Compatibilists resist the aspirations of libertarians to secure some form of absolute or ultimate agency (*qua* origination), they remain committed to the particular conception of responsibility encouraged by the

AU7



582 morality system and believe that it can be satisfied *within compatibilist*
583 *constraints* (see, e.g. Wallace 1994, 39, 64–66). It is these aims and aspi-
584 rations that Critical Compatibilism rejects. The trouble with Affirmation,
585 from this point of view, is that, on the face of it, it simply absorbs or
586 incorporates the orthodox Compatibilist aims and assumptions of the
587 morality system into its own proposed “solution” to the problem. If this
588 is the case, then both components or dimensions of Affirmation theory
589 retain the problematic baggage of “the morality system” and its peculiar
590 aims and assumptions.

591 The second point is intimately concerned with the first. A further
592 feature of “the morality system” is its aspiration to secure *optimism of a*
593 *certain kind*—a comforting and consoling picture about our ethical
594 lives and predicament. This feature of “the morality system” is deeply
595 resistant to any account that suggests that the exercise and operation of
596 our moral and rational capacities depends on large measure on factors
597 that are not controlled or governed by those same capacities and pow-
598 ers. All the parties in the freewill debate—libertarians, compatibilists,
599 and skeptics alike—hold onto EP and the particular form of optimism
600 that it insists on. This form of optimism denies the very *possibility* of
601 Free Will Pessimism, much less accepts it as the *truth* about our human
602 predicament.

603 We may, in light of the above, ask to what extent Affirmation remains
604 committed to this form of optimism as secured by an adherence and
605 commitment to EP? It is certainly evident that, unlike orthodox
606 Compatibilism or libertarianism, Affirmation dispenses with any pure or
607 simple optimism. It insists, more specifically, on the need to accommo-
608 date “dismay” as having a secure foothold, given the truth of Determinism.
609 However, as has also been explained, this form of dismay is grounded
610 entirely in skepticism and the origination of free will. Retaining this form
611 of dismay is consistent with remaining committed to an optimism that
612 rejects the very possibility of Free Will Pessimism (e.g. the skeptic may be
613 pessimistic because EP cannot be satisfied but will not allow that EP
614 should be discarded or abandoned altogether). As things stand, it is not
615 clear whether Affirmation is able or willing to accommodate forms of
616 dismay or pessimism based on rejecting EP and the associated aims and
617 aspirations of the morality system.

Let me conclude by describing an analogy that may help us to appreciate both the strengths and weaknesses of Affirmation. There is an aspect of Affirmation that we may think of in terms of the Duck/Rabbit Gestalt switch. It is a basic feature of Honderich's overall argument that a single image is not the whole truth on this topic. To insist on one side or the other as having the whole truth, and the other as being mistaken, is an error that both compatibilists and Incompatibilists share. We can and should, it is argued, accommodate and reconcile the partial truth contained in each perspective or position. We can, as it were, reconcile the view of the Incompatibilist Duck with that of the Compatibilist Rabbit, as long as we do not insist on the sole truth of either. Despite its attractions, however, this reconciliation project has its own vulnerabilities. More specifically, if Free Will Pessimism is right, both the Compatibilist and Incompatibilist components that are absorbed by Affirmation share a deeper and more problematic set of assumptions about the nature and conditions of freedom and moral responsibility. The aims and assumptions in question are those that are encouraged and endorsed by "the morality system." It may be that Affirmation can distance itself from these shared aims and assumptions but, as presented, it does not do this and, to this extent, it fails to identify the deeper difficulties that present themselves with respect to the consequences of Determinism and its relevance to the freewill problem as generally understood.⁴

Notes

1. My discussion in this chapter focuses largely on Honderich (2002), which, although it "follows the same path" as the earlier work, is intended to be more than a mere précis of it. As Honderich points out, *How Free Are You?* is not only shorter and more recent, it advances some new lines of argument and interest (2002, 6–7). For all these reasons, it seems the best work to focus my attention on for the purpose of this chapter—but readers should also consult Honderich's first and larger work for more detailed arguments and discussions relating to his views.
2. There are several overlapping features of Honderich's approach here that run parallel to P.F. Strawson's famous contribution in "Freedom

- 651 and Resentment” (1962). This includes the importance of not “over-
652 intellectualizing” this issue and, related to this, an understanding of
653 this debate in terms of broadly optimistic and pessimistic metaphysical
654 attitudes.
- 655 3. Whether we attach the label “fate” to this concern is merely a verbal mat-
656 ter—the substantial concern or issue remains with us.
- 657 4. It may be argued that Honderich’s commitment to a dualist picture of our
658 conceptual commitment and metaphysical attitudes requires a richer
659 genealogical approach—one that is more sensitive to how we (modern,
660 Westerners) have acquired the aims and assumptions of “the morality sys-
661 tem” and its specific views about moral freedom and ethical life. This is
662 certainly one feature that divides Honderich’s (ahistorical) approach from
663 William’s more self-conscious genealogical account. See, in particular,
664 Williams (1993).

665 References

[AU8](#)

- 666 Ayer, A. J. 1954. Freedom and necessity. Reprinted in *Free will*, ed. Gary Watson,
667 15–23. Oxford: Oxford University Press.
- 668 Caruso, Gregg D. forthcoming. Free will skepticism and its implications: An
669 argument for optimism. In *Free will skepticism in law and society*, ed. Elizabeth
670 Shaw and Derk Pereboom. Cambridge, MA: Cambridge University Press.
- 671 Dennett, Daniel. 1984. *Elbow room: The varieties of free will worth wanting*.
672 Oxford: Clarendon Press.
- 673 Honderich, Ted. 1988. *A theory of determinism: The mind, neuroscience and life*
674 *hopes*. Oxford: Oxford University Press.
- 675 ———. 2002. *How free are you?* Oxford: Oxford University Press.
- 676 Nagel, Thomas. 1976. Moral luck. Reprinted in *The philosophy of free will:*
677 *Essential readings from the contemporary debates*, eds. Paul Russell and Oisín
678 Deery, 31–42. New York: Oxford University Press.
- 679 Pereboom, Derk. 2001. *Living without free will*. Cambridge, MA: Cambridge
680 University Press.
- 681 ———. 2013. Optimistic skepticism about free will. In *The philosophy of free*
682 *will: Essential readings from the contemporary debates*, ed. Paul Russell and
683 Oisín Deery, 421–449. New York: Oxford University Press.
- 684 ———. 2014. *Free will, agency, and meaning in life*. Oxford: Oxford University
685 Press.

9 Free Will and Affirmation: Assessing Honderich's Third Way

- Russell, Paul. 2000. Compatibilist fatalism. Reprinted in *The philosophy of free will: Essential readings from the contemporary debates*, eds. Paul Russell and Oisín Deery, 450–468. New York: Oxford University Press. 686–688
- . 2002. Pessimists, pollyannas and the new compatibilism. In *The Oxford handbook of free will*, ed. Robert Kane, 229–256. New York: Oxford University Press. 689–691
- . 2017a. Free will pessimism. In *Oxford studies in agency and responsibility*, vol. 4, ed. David Shoemaker. Oxford: Oxford University Press. Reprinted in Russell (2017b). 692–693
- . 2017b. *The limits of free will: Selected essays*. New York: Oxford University Press. 694–696
- Russell, Paul, and Oisín Deery, eds. 2013. *The philosophy of free will: Essential readings from the contemporary debates*. New York: Oxford University Press. 697–698
- Smilansky, Saul. 2000. *Free will and illusion*. Oxford: Oxford University Press. 699
- Strawson, Peter F. 1962. Freedom and resentment. Reprinted in *The Philosophy of free will: Essential readings from the contemporary debates*, eds. Paul Russell and Oisín Deery, 63–83. New York: Oxford University Press. 700–702
- Strawson, Galen. 1994. The impossibility of ultimate moral responsibility. Reprinted in *The philosophy of free will: Essential readings from the contemporary debates*, eds. Paul Russell and Oisín Deery, 363–378. New York: Oxford University Press. 703–706
- Wallace, R. Jay. 1994. *Responsibility and the moral sentiments*. Cambridge, MA: Harvard University Press. 707
- Waller, Bruce. 2011. *Against moral responsibility*. Cambridge, MA: MIT Press. 709
- Watson, Gary, ed. 1982. *Free will*. Oxford: Oxford University Press. 710
- Williams, Bernard. 1985a. *Ethics and the limits of philosophy*. With a commentary by A.W. Moore and a foreword by Jonathan Lear. Routledge: London and New York: Routledge Classics (edition 2011). 711–713
- . 1985b. How free does the will need to be?. Reprinted *Making sense of humanity*, ed. B. Williams, 3–21. Cambridge: Cambridge University Press (1995). 714–716
- . 1993. *Shame and necessity*. Berkeley: University of California Press. 717

Author Queries

Chapter No.: 9 0003197158

Queries	Details Required	Author's Response
AU1	Please check and confirm if the affiliations are presented correctly.	
AU2	Please note that the Roman numbers "I" and "II" have been set as section headings. Kindly check.	
AU3	In the part "determinism comes to no more that a yes answer to the question" in the extract, should "that" be changed to "than"? Please check the original and confirm.	
AU4	Is uppercase needed for "Free Will Pessimism"?	
AU5	Please check if the expression " β . < Exclusion Premise (EP) >." Is complete.	
AU6	Please check whether it is okay to delete this dash, below Point 4.	
AU7	Please specify "a" and "b" for the reference "Williams (1985)".	
AU8	References "Russell (2002, 2017b), Russell and Deery (2013), and Watson (1982)" were not cited anywhere in the text. Please provide a citation.	



Uncorrected Proof