*COMPATIBILIST-FATALISM**

Paul Russell
University of British Columbia

Published in Ton van den Beld, editor,
Moral Responsibility and Ontology (Kluwer: Dordrecht, 2000), 199-218.

Compatibilists argue, famously, that it is a simple incompatibilist confusion to suppose that determinism implies fatalism. Incompatibilists argue, on the contrary, that determinism implies fatalism, and thus cannot be consistent with the necessary conditions of moral responsibility. Despite their differences, however, both parties are agreed on one important matter: the refutation of fatalism is essential to the success of the compatibilist strategy. In this paper I argue that compatibilism requires a richer conception of fatalistic concern; one that recognizes the *legitimacy* of (pessimistic) concerns about the origination of character and conduct. On this basis I argue that any plausible compatibilist position must concede that determinism has fatalistic implications of some significant and relevant kind, and thus must allow that agents may be legitimately held responsible in circumstances where they are subject to fate. The position generated by these compatibilist concessions to incompatibilism will be called 'compatibilist-fatalism'.

I

Compatibilist-fatalism has two key components:

(1) It claims that the truth of determinism is compatible with conditions of responsibility. I will call this the 'responsibility-compatibilist claim'. (Its contrary will be called the 'responsibility-incompatibilist claim'.)
(2) It claims that determinism implies conditions of universal fatalism. I will call this the 'fatalism claim'.
There is near unanimous agreement in both compatibilist and incompatibilist camps that it is incoherent to combine these two claims, since an agent cannot be both responsible and subject to fate. Compatibilists and incompatibilists have, nevertheless, very different reasons for taking this view. Indeed, their superficial agreement conceals fundamental differences about the nature and significance of *fatalism* itself.

With remarkable consistency compatibilists have been very clear about why they believe that the fatalistic claim should be rejected. It is, they maintain, a product of simple confusion - a confusion that gives illegitimate support to incompatibilism. The compatibilist argument against the fatalism claim - let us call it the 'refutation argument' - is very familiar. In an influential statement of classical compatibilism R.E. Hobart gives the following brief account of the refutation argument:

> Fatalism says that my morrow is determined no matter how I struggle. This is of course a superstition. Determinism says that my morrow is determined through my struggle ... The stream of causation runs through my deliberations and decisions, and if it did not run as it does run, the event would be different.[1]

According to this view, then, determinism is the thesis that everything that occurs, including our deliberations and decisions, are causally *necessitated* by antecedent conditions. Fatalism, by contrast, is the doctrine that our deliberations and decisions are causally *ineffective* and make no difference to the course of events. In circumstances of fatalism what happens does not depend on how the agent deliberates. The relevant outcome will occur no matter what the agent decides. Clearly, however, determinism does not imply fatalism. While there are some circumstances in which deliberation is futile (i.e. 'local fatalism'), deliberation is nevertheless generally effective in a deterministic world.[2]

Let us call those who accept the responsibility-compatibilist claim but reject the fatalist claim 'orthodox-compatibilists'. Orthodox-compatibilist understanding of the relationship between responsibility and fate seems clear enough - indeed, one of its attractions is its simplicity. In circumstances where a person is subject to fate, her deliberations and decisions cannot change the course of events. Whatever occurs in these circumstances does not depend on the agent's deliberations. Accordingly, if it were true that determinism implied universal fatalism then it would follow that no one would be responsible - since no one would be able to influence or alter what occurs.[3] However, as the refutation argument makes plain, none of these consequences follow from the truth of determinism. Responsibility-incompatibilism, therefore, has no legitimate foundation in the fatalism claim that incompatibilists mistakenly try to draw from the thesis of determinism.

Incompatibilists defend the fatalism claim and reject the (orthodox) compatibilist's refutation-argument.[4] The incompatibilist reply to the refutation argument turns, crucially, on an alternative interpretation of *fate*. The incompatibilist maintains that compatibilist accounts of 'fate', interpreted in terms of the causal ineffectiveness of an agent's deliberations and actions, is wholly inadequate, and that in consequence it evades not only real difficulties of a fatalistic character, but also related difficulties about the conditions of moral responsibility. Incompatibilist concern about fate is not - as on the refutation argument - directed to the issue of the causal influence *of the agent*, but rather at the issue concerning the causal influence on the agent. An agent is said to be subject to 'fate', on this account, if her character and conduct does not (ultimately) originate with the person concerned. The incompatibilist maintains that determinism implies universal fatalism in the sense that - however complex the mechanisms at work may be - the causal chains eventually reach outside the agent, and hence no person is *the real originator or ultimate source* of her conduct and character. When an agent is not the (ultimate) source of her actions then, the incompatibilist argues, the person is subject to fate.

Compatibilists and incompatibilists, evidently, conceive of 'fate' in quite different terms. For the compatibilist a person is subject to fate only if their circumstances are such that they are unable to causally *contribute* to the course of events in some relevant respect. Let us call this account of fate, as developed in the refutation argument, the concept of 'contributory-fate'. Incompatibilists do not (or need not) deny that contributory-fate is one mode of fatalistic concern, nor need they suppose that determinism implies that contributory-fatalism holds universally.[5] What the incompatibilist maintains is that there is another mode of fatalistic concern that arises from a backward-looking perspective (and is, as I will explain, intimately linked with problems of responsibility). The question that concerns us from this perspective is whether or not the agent is the ultimate source or true originator of her character and conduct. An agent is subject to 'fate', in this sense, if her circumstances are such that her character and conduct have origins and sources that (ultimately) extend beyond her.[6] Let us call this alternative, incompatibilist conception of fate 'origination-fate'. The essence of the incompatibilist position is that determinism implies that origination-fate is the universal condition, and thus renders responsibility impossible.[7]

Issues of responsibility and fatalism are intimately and inextricably woven-together on the standard incompatibilist account. Incompatibilists object to the compatibilist's refutation argument on the ground that it constitutes a superficial response on this issue (i.e. fate), and argue that it reflects a one-sided, forward-looking pragmatic perspective that fails to capture - or even acknowledge - difficulties arising from the backward-looking perspective (i.e. matters of origination as opposed to contribution). According to the incompatibilist, the very same shortcomings can be found in compatibilist views on responsibility, and for reasons that are rooted in and run parallel to the failings of the refutation argument.[8] Incompatibilists grant that it is

possible to advance a 'superficial' conception of responsibility that is essentially pragmatic and forward-looking in nature, and this can be reconciled with determinism. What cannot be reconciled with determinism, however, is deep responsibility.[9] Deep responsibility is concerned not with the causal efficacy of the attitudes and practices of blaming and punishing, but rather with whether these attitudes and practices are deserved or merited. To understand (deep) responsibility in these terms involves a change of perspective from forward-looking to backward-looking considerations. Only from this perspective can we understand the retributive aspects of responsibility which the compatibilist's (superficial) forward-looking account cannot capture.

It is at this point that incompatibilists draw on their defence of the fatalism claim, and use it to support their responsibility-incompatibilist conclusion. Attributions of desert, claims the incompatibilist, rest with an agent's capacity for *self-determination*, and this requires the metaphysics of indeterminism. The incompatibilist maintains, in other words, that it will not suffice to establish a person's responsibility to show, simply, that her deliberations and conduct are causally effective in the world. On the contrary, what is required is to show that the choices and actions originate with the agent - and that is why we hold the agent accountable. Clearly, then, since determinism implies universal origination-fatalism, it makes responsibility impossible. It is in this manner that the responsibility-incompatibilist claim and fatalism claim are inextricably bound together on the standard incompatibilist account.[10]

II

The success of any compatibilist strategy depends on showing that 'origination', understood in terms of indeterministic metaphysics, is not a necessary condition of moral responsibility, and that a suitably 'deep' account of responsibility can be provided within the restrictions imposed by compatibilist commitments. It is not possible in this context to provide any full-scale defence of the case for responsibility-compatibilism. For our purposes, however, this is not necessary. All that is necessary is to describe the general structural features of the case for responsibility-compatibilism in order to assess its significance for the *distinct* prospects of compatibilist-fatalism (as contrasted with orthodox-compatibilism). Suffice it to say, that if there is nothing of a convincing nature to be said in support of the responsibility-compatibilist claim then both orthodox-compatibilism *and* compatibilist-fatalism collapse - since this claim is common to both.

There are two independent but merging strands in contemporary compatibilist thinking that promise a 'deeper' and more 'robust' compatibilist account of moral responsibility. Both these strands can be described under the general heading of 'naturalized responsibility'. The first strand is closely associated, in the contemporary context, with P.F. Strawson's highly influential paper 'Freedom and Resentment'.[11] The view advanced is that circumstances of responsibility must be understood in terms of the natural workings of moral sentiment. Human beings, it is argued, are inescapably subject to moral emotions under certain conditions, and no general 'theoretical' considerations concerning the truth of determinism can discredit - much less dislodge - our human commitment to these emotional responses. To suppose otherwise is to 'over-intellectualize' these matters. The most notable strength of this strand of naturalized responsibility is that it distances itself from the cruder utilitarian, forward-looking features of classical compatibilism, without making any concessions to the metaphysics of libertarianism. The Strawsonian strategy, therefore, plugs a significant 'gap' in the compatibilist position, and provides a substantial basis for accounting for backward-looking, desert-based considerations consistent with compatibilist commitments. To

this extent, compatibilists are better placed to provide their account of responsibility with the kind of 'depth' which it plainly requires.

Although the Strawsonian strand of naturalized responsibility (plausibly) addresses a number of traditional incompatibilist objectives, it has its own significant vulnerabilities. The most important of these is, perhaps, that taken by itself it fails to explain on what basis individuals are or are not appropriate objects of moral sentiment.[12] More specifically, without some account of the relevant *capacities* required of moral agents, the theory remains entirely open to the incompatibilist counter-argument that what is required is some mode of contra-causal freedom. The second strand of contemporary naturalized responsibility, however, appears to plug this gap in the position very neatly. What is required is an account of moral capacity that can account for freedom of the *will*, as well as freedom of action. Various models of 'hierarchical' or 'real self' theories provide this.[13] Moral freedom, it is argued, is not simply a matter of being able to act according to your own will, unimpeded by external constraints. It also involves a capacity to reflect on the structure of your own will and form preferences about which desires move you to action. On the basis of a higher-order capacity of this kind agents are able to 'identify' with or 'repudiate' their own will - something that is essential to being capable of moral conduct and an appropriate object of moral sentiment. It is a general capacity of this nature that distinguishes fully responsible human adults from animals and children who (in some degree) do not enjoy such a capacity and thus are not (fully) responsible. The crucial point remains, however, that this sort of higher-capacity involves no contra-causal or libertarian metaphysical commitments.

Although it would be entirely premature to declare this two-pronged defence of the responsibility-compatibilist claim a success (as clearly the matters raised continue to be strenuously debated), it is nevertheless fair to say that this general approach provides substantial support for the position taken.[14] Let us say, therefore, that the responsibility-compatibilist claim has substantial (although not conclusive) support. The issue that concerns us is what the implications of this are for the compatibilist position in respect of the matter of fatalism. The view that is most widely accepted on this issue is plain. If responsibility-compatibilism is accepted, then the fatalism claim must be rejected, as both cannot be accepted.

Let us call the assumption that responsibility and fate *exclude* each other the 'exclusion thesis'. Both orthodox-compatibilists and incompatibilists accept the exclusion thesis although, as I have explained, they accept it for very different reasons. The exclusion thesis, however, provides a very quick way of dealing with the issue of fatalism once the responsibility-compatibilist claim is established. The exclusion thesis eliminates the possibility that conditions of universal fatalism could persist in conditions when agents are still morally responsible. Hence, if agents are responsible, conditions of universal fatalism cannot hold. In short, if we accept the responsibility-compatibilist claim, and the exclusion thesis, then we *must* reject the fatalism claim. If this is correct, then compatibilist-fatalism is an untenable position.

If a case can be made for compatibilist-fatalism it must be able to show that there is some basis for accepting the fatalism claim without compromising the responsibility-compatibilist claim (thereby showing the exclusion thesis to be false). Another way of expressing this is to say that there must be issues of fatalism that survive the (assumed) success of responsibility-compatibilism. On the face of it, however, this is odd, as incompatibilist concern about the fatalistic implications of determinism (i.e. in respect of origination-fate) are generally motivated by worries about responsibility-incompatibilism. The puzzle is, that if determinism has no responsibility-incompatibilist implications then the issue of origination-fatalism seems to be empty.[15]

III

In order to explain the distinctive commitments of compatibilist-fatalism it will be useful to employ the terminology of 'optimism' and 'pessimism'.[16] These labels are illuminating for understanding the free will debate because they indicate that the various parties involved have certain concerns or interests that motivate the positions that they take. In other words, these labels make plain that the issues at stake are not merely theoretical (conceptual) puzzles that require clarification but, rather, they are matters that are in some sense emotionally charged. The language of 'pessimism', in particular, is indicative of the fact that incompatibilists find some implications of determinism troubling or disturbing.[17] For the incompatibilist determinism suggests a picture of human beings that is (somehow) disillusioning, and thus the incompatibilist wants this thesis to be false. Compatibilists, by contrast, do not share these concerns, and believe, indeed, that they are misguided and a product of (philosophical) confusion. Since compatibilists find nothing 'troubling' or 'disturbing' about the thesis of determinism - and nothing about it motivates a desire that it be false - they may be characterized as 'optimists'.

Any position that accepts the fatalism claim seems to be committed to pessimistic motivations of some kind. In the case of incompatibilism these pessimistic motivations, as we have noted, are closely tied to concerns about the conditions of responsibility. These concerns are not endorsed by compatibilist-fatalists since they accept the (contrary) responsibility-compatibilist claim. The obvious question arises, therefore, given their commitment to the fatalism claim, what are the pessimistic motivations of the compatibilist-fatalist? Clearly compatibilist-fatalists hold that determinism implies universal origination-fatalism and there is something 'troubling' or 'disturbing' about this which lies beyond the scope of issues of responsibility. However, the source of this pessimism remains obscure.

What is essential to compatibilist-fatalism is the view that while origination-fatalism does not undermine or discredit our (natural) commitment to moral responsibility, it nevertheless does not leave our conception of ourselves as real agents in the world undiminished. A well-known passage of Spinoza's *Ethics* identifies this source of pessimistic concern and describes it in the following terms:

> Most of those who have written about the emotions and human conduct seem to be dealing not with natural phenomena that follow the common laws of Nature but with phenomena outside Nature. They appear to go so far as to conceive man in Nature as a kingdom within a kingdom. They believe that he disturbs rather than follows Nature's order, and has absolute power over his actions, and is determined by no other source than himself.[18]

Spinoza's observations appear in a context in which he is seeking to explain the source of deep *resistance* to any naturalized, deterministic conception of human life. Although much of this resistance is motivated by incompatibilist concerns about the threat to the fabric of moral responsibility, Spinoza's remarks bypass them. Instead, his remarks are addressed directly at the issue of agency. The specific dimension of pessimistic concern is captured through the metaphor of 'sovereignty'. In conceiving of human beings as 'a kingdom within a kingdom' we conceive of ourselves as subject, not to the alien laws that govern all nature, but rather to laws that pertain uniquely to human (rational) life. Our sense of 'sovereignty', therefore, is tied to our belief that we

are distinct from nature, not (a reducible) part of it. Through our capacity for sovereignty, so conceived, we are not only independent of nature, but also *above* it. We are above it - qua sovereign - because we govern nature without being governed by it (i.e. we are not subject to its laws).[19] From this perspective we take ourselves to be something more than (sophisticated and complex) causal intermediaries. We conceive of ourselves as starting points that intervene in the order of things. Finally, the metaphor of sovereignty brings with it a conception of beings who are worthy of a particular kind of dignity - the dignity due to beings who are sovereign over both themselves and nature.

Clearly, from this perspective, we want much more than simply to be morally accountable to each other. What is at stake here is our conception of ourselves as (actively) *ordering nature*, rather than being (passively) *ordered by nature*. This distinction depends on a capacity for spontaneous self-determination and thus cannot be sustained in conditions of universal origination-fatalism. Any optimism secured on the basis of responsibility-compatibilism, therefore, must be significantly tempered by a pessimism rooted in these reflections. Although we may concede that universal origination-fatalism poses no threat to the fabric of responsibility, it nevertheless has troubling implications for aspects of our self-conception that lie outside this sphere (something that is obscured by incompatibilist arguments that focus exclusively on issues of responsibility). Only those who are unmoved by the issue of 'sovereignty', and place no value on it, can draw any other conclusion.

What reply can orthodox-compatibilists offer to this line of reasoning? The first point to note is that it will not do to fall back on the refutation argument. The pessimistic concerns of the compatibilist-fatalist are not motivated by any simple confusion between determinism and contributory-fatalism. On the contrary, compatibilist-fatalists (along with incompatibilists) object to the refutation argument on the ground that it fails to draw the relevant distinction between origination and contributory fate and is, consequently, blind to the very different concerns that arise from the former. Furthermore, the entire line of reasoning that develops from the refutation argument proceeds from the same one-sided, forward-looking perspective that generated serious shortcomings in the efforts of classical compatibilists to address incompatibilist concerns about responsibility. Since compatibilism has overcome its blindness to backward-looking claims in respect of responsibility, so too it must face the issues raised by origination-fatalism in a more direct manner.

The orthodox-compatibilist may argue that it is possible to address these concerns about origination without accepting the fatalism claim. It may be argued, for example, that the resources of naturalized responsibility provide an effective basis from which to discredit the specific concerns that the compatibilist-fatalist has raised. What is supposed to be troubling about determinism is that it makes genuine origination or (true) self-determination impossible. If there is any foundation to the pessimistic concerns that support the fatalism claim, this seems to be it. Against this, however, it can be argued that hierarchical or real self theories of freedom provide a substantial account of self-determination and self-control without any appeal to indeterministic metaphysics. All that is required is a suitably complex description of the higher-order capacities of human beings to reflect on their own character and motivation and restructure their own wills on this basis. It is simply incorrect, on this account, to suppose that any agent in a deterministic framework is incapable of altering or amending his character and the structure of his own will. Agents with the relevant capacities of the sort described (i.e. two-level freedom) are not passive in these respects. Indeed, with capacities of these (natural) kinds we can, to a large extent, conceive of ourselves as 'self-

made-selves'.[20] Whatever residue of pessimistic concern survives responsibility-compatibilism, therefore, is effectively discredited by these considerations.

Does this orthodox-compatibilist counter-argument - let us call it the '*revised* refutation argument' - serve to discredit the distinct pessimistic concerns that motivate the compatibilist-fatalist? The revised refutation argument is obviously an improvement on the original argument. It does not, for example, suggest that the defender of the fatalism claim makes the crude mistake of supposing that determinism implies universal *contributory*-fatalism. More importantly, this revised effort to refute the fatalism claim does not deny the general legitimacy of concerns that arise from a backward-looking perspective about the issue of origination of character and conduct. What is argued is that determinism provides no basis for pessimistic concerns of this kind and to this extent the concerns are unreasonable. The resources of higher-order capacities are more than adequate to account for talk of self-determination and self-control (i.e. some form of 'sovereignty') and they do so without relying on the obscure metaphysics of libertarianism to fill this particular gap.

The strength of the revised argument is that it shows that compatibilists can provide a more sophisticated account of self-determination and freedom of will, which is a clear improvement on the more limited (classical) compatibilist accounts of freedom understood in terms of unimpeded action. Nevertheless, it is not evident that the case against the fatalism claim can be secured by means of the revised refutation argument. The (higher-order) moral capacities described may well serve as the relevant basis on which to distinguish individuals who are appropriate objects of moral sentiment from those who are not. (Indeed, for reasons that have been explained, the case for responsibility-compatibilism depends on this.) However, capacities of these kinds are not capable of addressing the specific difficulties that are suggested by reflection on the implications of (universal) origination-fatalism.

First, the compatibilist-fatalist may grant that human beings have capacities of self-determinism of the sort described without in any way conceding that these capacities are of such a nature as to allow agents to reinvent themselves as they please. Any account of these capacities, so construed, is self-evidently an exaggeration. Clearly there are many other forces of an external nature that condition character and the conduct that flows from it. Accordingly, the scope and extent of the human capacity for self-determination of this sort is much more limited and restricted than orthodox-compatibilist talk of 'self-made-selves' suggests.[21] Second, and more importantly, even if these powers of self-control were as extensive as defenders of the revised refutation argument imply, they entirely fail to address the more basic concern that sustains the fatalism claim. The specific concern is that ultimately nothing that the agent is or does originates with the agent - the causal source can always be traced to factors lying outside the agent. Granted a deterministic framework, when and how an agent actually *exercises* such capacities of rational self-criticism and redirection will depend, ultimately, on factors that lie beyond the agent.[22] This brings us back full-circle to the specific implication of determinism that compatibilist-fatalists find disturbing: determinism implies that no agent is the ultimate source of her own character and conduct.[23]

This basic concern is, of course, very familiar in literature critical of compatibilist efforts to account for self-determination. It is, however, particularly important to note that while libertarian efforts to explain what ultimate agency consists in may be judged hopelessly obscure, the aspiration itself is motivated by a general worry that is clear enough: namely, that compatibilist accounts of self-determination are essentially superficial, since such agents are, inescapably, conditioned by factors that they have no control over. Clearly, then, the revised refutation argument, fails to discredit this fundamental concern. It may be argued, furthermore, that this

conclusion is especially disturbing if the compatibilist is right, and our natural commitment to responsibility persists in face of these (fatalistic) conditions.[24]

IV

In face of this reply to the revised refutation argument, orthodox-compatibilists may suggest another way of discrediting the pessimistic concerns that seem to sustain the fatalism claim. What is not clear, they may argue, is what sort of 'origination' or 'self-determination' is required to avoid these fatalistic anxieties. More specifically, the desire to be a (pure) self-determinator, so conceived, is simply incoherent, and thus no real sense can be made of the pessimistic concerns that lie behind the fatalism claim.[25] Moreover, in so far as any sense can be made of this desire for (pure) self-determination it appears, on examination, less than desirable. So the orthodox-compatibilist reply is this: the objective of 'overcoming' origination-fate in the terms suggested is neither coherent, nor obviously attractive in itself. To this the orthodox-compatibilist may also add that it is important to note that the problem of fate, conceived in terms of worries about origination (rather than contribution), is not limited to the metaphysics of determinism. On the contrary, the metaphysics of indeterminism generates its own 'fatalistic' worries in this regard. That is, even if there are real 'breaks in the causal chain', and 'spontaneous willings' occur, it is not evident that this serves to secure 'genuine agency'. This is because (pure) 'spontaneity' seems to undermine genuine agency no less than the chains of causal necessity. The underlying point is, of course, that the ideal of 'genuine agency' is simply a confused illusion that cannot stand up to critical scrutiny. Given this, the pessimistic concerns that are supposed to sustain the fatalism claim can be dismissed as wholly unreasonable.

This rejoinder seeks to discredit the pessimistic motivations of the compatibilist-fatalist by arguing that there is no plausible *alternative* metaphysics that could overcome these difficulties (i.e. regarding ultimate self-determination or origination). In my view, however, this is not a convincing way to discredit these concerns about origination-fatalism. The obvious point is that it may be granted that there is no alternative metaphysics that serves to insulate us from these pessimistic concerns about the ultimate origination of character and conduct, but this does not show that these concerns are somehow bogus or without foundation. Consider, for example, the analogous debate concerning the doctrine of the immortality of the soul. Many philosophers - especially religiously-minded philosophers - have argued that we have reason to want to be immortal, to exist for all eternity. Accordingly, faced with arguments for human mortality (i.e. naturalistic conceptions of human beings) these philosophers maintain that mortalism has pessimistic implications. Against pessimism of this nature, defenders of mortalism may argue (in parallel reasoning with orthodox-compatibilists) that the desire for immortality is neither coherent in itself, nor an obviously attractive ideal - to the extent that we can conceive of it being realized.

Clearly, however, those who find mortalism a source of pessimism (i.e. troubling; difficult; disillusioning; and so on) may readily grant the truth of the mortalist's claims concerning the doctrine of immortality. Nevertheless, it simply does not follow that if one grants that the desire for immortality involves an ideal or aspiration that is doubtfully coherent and (on reflection) doubtfully attractive, then there is no basis for being troubled by reflections on human mortality.[26] On the contrary, reflection on this specific aspect of the human condition provides a reasonable basis for being troubled *whatever view we take*. There is no guarantee that some trouble-free optimistic alternative must be 'available' to us. Indeed, in the case of human mortality/immortality the truth

seems otherwise. What is troubling about human mortality is that it confronts us with the limits of human existence - our inevitable and inescapable *finitude* as beings in the world. Clearly, then, while we may not want to be immortal, and we may agree there is no coherent account of what we would want if we sought immortality, we may still have some reasonable basis for finding the limits of human existence and individual finitude matters that unsettle and disturb us in important respects (so long as we are tolerably reflective on the matter). This feature of the human condition is something that we cannot contemplate with optimistic calm and serenity.[27]

Parallel reasoning is available to the compatibilist-fatalist. Against this position it is argued that concerns about the fatalistic implications of determinism rely upon an ideal of (pure) self-determination that is neither coherent nor, on reflection, attractive. From this the orthodox-compatibilist concludes that there is no basis for the pessimistic anxieties that are supposed to sustain the fatalism claim. To this, however, the compatibilist-fatalist may reply that, however incoherent and unattractive the ideal of pure (unconditioned) agency may be, what is troubling about origination-fatalism is that it confronts us with the limits of human agency - the inescapable fact that the ultimate source of our character and conduct lies beyond us.[28] Our finitude and place in the order of nature has implications for our conception of ourselves as genuine agents. We may not want to be (God-like) self-creators, and we may agree that there is no available coherent interpretation of this ideal, and yet consistently maintain that reflections on these limits concerning the origination of human agency are disturbing and troubling in ways that are analogous to reflections about human mortality. To insist on (easy) optimism in face of such thoughts about the human condition is a form of 'superficiality' to which (orthodox) compatibilists are much too prone.

V

There is one final reply to the compatibilist-fatalist that may now be presented. The fatalism claim receives whatever support it has on the basis of the pessimistic concerns that it generates from reflections about origination. The orthodox-compatibilist may simply insist that none of these concerns move him, or *trouble* him, in the least. It may be argued, moreover, that it is the compatibilist-fatalist who is guilty of 'over-intellectualizing' this whole issue by appealing to 'theoretical' considerations regarding origination in order to *compel* a particular affective response (i.e. pessimism) - but this cannot be done.

This reply, however, is one that the orthodox-compatibilist should be reluctant to employ. The orthodox-compatibilist has tried to *discredit* the fatalism claim by showing that, in some way or other, it depends on confusion and/or illusion. In reply it has been shown that these attempts to refute the fatalism claim are themselves confused, or manifest a shallow appreciation of fatalistic concern. The compatibilist-fatalist may grant, at this stage, that their concerns may not be shared by everyone, and that it is impossible to argue someone into the relevant attitude (i.e. pessimism) once all relevant considerations have been made clear. Nevertheless, if it is impossible to compel pessimistic attitudes in face of such considerations, it is no less impossible to compel optimism. As there seems to be no identifiable confusion lying behind either the optimistic or pessimistic attitude in these circumstances, a stalemate results. This situation, however, leaves orthodox-compatibilists unable to discredit the pessimism that sustains the fatalism claim. All that can be said in reply is that the orthodox-compatibilist does not share it, which is clearly a different matter. It suffices, therefore, that the pessimism that motivates the commitment to the fatalism claim has not been discredited, and the orthodox-compatibilist is mistaken to suppose that it can be.

The compatibilist-fatalist may also argue that the best explanation for the fact that orthodox-compatibilists are unable to share this pessimism is that they have not sufficiently exercised their reflective imagination. To remedy this, they may suggest that appropriate reflection on especially striking cases will help to make clear why pessimistic concerns about origination are called for.[29] As I have explained, however, it would be a mistake to represent the pessimistic concerns that sustain commitment to the fatalism claim as simply the end-result of a process of pure reasoning, as clearly such concerns also require some relevant *sensibility*. (Consider, again, the analogy with pessimistic reflections on death.) This is why the cultivation of artistic imagination is of such obvious significance in this context; since many great works of literature and drama are devoted to the central message of compatibilist-fatalism (i.e. that responsibility and fate come fused together in human life).[30]

Another possibility is to show that the orthodox-compatibilist's inability to share this mode of pessimism is rooted in confusion about the *quality* of the pessimism involved. Pessimism varies in its quality as well as its source. The quality of pessimism generated by contributory-fatalism may be characterized as one of *despair*, produced by a sense of impotence. To conceive of ourselves as 'puppets' or 'dolls', for example, would certainly be awful and justify despair.[31] The pessimism associated with origination-fatalism, however, is not of this character.

Origination-fatalism, I have argued, focuses on our awareness of human finitude and its relevance to agency. This basic concern is well captured by John Macquarrie in the context of a discussion of existentialist philosophy.

> Man is thrown into existence, each one is thrown into his own particular existential situation. From the human point of view, it is rather like the throw of a dice.... As we see it from the purely human point of view, we all start out as different people with different endowments in different situations, and there is as little assignable reason for the differences as there is for the dice turning up one number rather than another.[32]

As these remarks suggest, the pessimism of compatibilist-fatalism is not so much a sense of despair rooted in impotence, but rather one of being *disconcerted*, rooted in awareness of finitude and contingency.[33] Closely associated with the sense of finitude and contingency is, I suggest, a sense of the absurdity of human life.[34] In this context it takes the form of an (uncomfortable) awareness of the gap between our aspiration to 'sovereignty' and being 'self-made-selves', and the recognition, as conveyed by the fatalism claim, that this is an illusion. It is evident, therefore, that the pessimism involved in endorsing the fatalism claim, so interpreted, is of a very different kind than the pessimism associated with contributory-fatalism (i.e. as featured in the refutation argument). Much of the orthodox-compatibilist resistance to the pessimism of compatibilist-fatalism is based, I suggest, in a confusion between these two very different modes of fatalistic concern, and the distinct sensibilities associated with them.

It should now be clear what the optimistic and pessimistic commitments of compatibilist-fatalism come to. In respect of the issue of responsibility, the compatibilist-fatalist maintains that the resources of naturalized responsibility are rich enough to provide firm support for the responsibility-compatibilist claim. (This is an issue that I have left open, except for the proviso that a strong enough case can be made for this claim to give it considerable credibility.) In respect of the fatalism claim compatibilist-fatalists hold that the refutation argument is blind to pessimistic concerns about origination. Moreover, even the more substantial revised version of the

refutation argument (employing the resources of naturalized responsibility) cannot discredit or dislodge the source of pessimism that sustains commitment to the fatalism claim. So interpreted, compatibilist-fatalism involves *mixed* optimistic and pessimistic elements, and to this extent it addresses both compatibilist and incompatibilist concerns.

VI

My objective in this paper has not been to defend compatibilist-fatalism, but to consider its merits in relation to orthodox (non-fatalistic) compatibilism. Both forms of compatibilism accept the responsibility-compatibilist claim: that is, the claim that determinism does not discredit the attitudes and practices associated with moral responsibility. Where they diverge is on the matter of fatalism. Compatibilist-fatalists accept a claim that is generally associated with incompatibilism: namely, that determinism has fatalistic implications. The discussion in this paper, therefore, has been primarily concerned to provide and an interpretation and defence of the fatalism claim from the perspective of those who are already (i.e. independently) committed to the responsibility-compatibilist claim. For reasons that have been explained, this is an unusual and controversial position for any compatibilist to adopt.[35]

I have described a number of different approaches that the orthodox-compatibilist may take in order to discredit the specific pessimistic motivations associated with the fatalism claim. All of them, I argue, are unsuccessful. It follows from this that any plausible compatibilism must take the form of - or accept the legitimacy of - compatibilist-fatalism.36 An obvious corollary of this is that a plausible compatibilism must reject the exclusion thesis.[37] A particular merit of compatibilist-fatalism is that it recognizes the (deep) source of incompatibilist intuitions as rooted in backward-looking concerns about the origination of character and conduct and, related to this, it avoids the one-sided superficiality of the (classical) refutation argument. When these points are properly established, I maintain, the compatibilist is better placed to provide a more nuanced and appropriate response to the (pessimistic) concerns of the incompatibilist

A plausible compatibilism, I conclude, must embrace a richer conception of fatalistic concern, and allow for the possibility that agents may be legitimately held responsible in circumstances where they are subject to fate. The significance of this should be clear. Hitherto all forms of compatibilism have been orthodox in character: they reject the fatalism claim and are homogeneously 'optimistic'. The central thesis of this paper is that compatibilism can (or must) take the form of compatibilist-fatalism, and thereby accept that determinism has fatalistic implications without compromising its commitment to naturalized responsibility.

FOOTNOTES

* I am grateful to Ton van den Beld, Richard Double, Richard Gale, Walter Glannon, David Gauthier, Saul Smilansky, Jay Wallace, and Allen Wood for helpful comments and suggestions on earlier drafts of this paper. For further helpful comments and discussion I would like to thank audiences at Simon Fraser, Washington (Seattle), British Columbia, Edinboro (Pennsylvania), Virginia and, especially, the Utrecht conference on 'Moral Responsibility and Ontology' (1998).

(1) R.E. Hobart, 'Free Will as involving Determination and Inconceivable Without It', reprinted in Bernard Berofsky, ed, *Free Will and Determinism* (New York: Harpers & Row, 1966), 82.

(2) Daniel Dennett is the most prominent contemporary defender of the (classical) refutation argument. As an example of 'local fatalism' he describes circumstances where a person has thrown himself off the Golden Gate Bridge and then asks if this is really such a good idea. For this person, Dennett observes, 'deliberation has indeed become impotent'. Dennett, *Elbow Room: The Varieties of Free Will Worth Wanting* (Oxford: Clarendon Press, 1984), 104. The point is, however, that these circumstances are 'abnormal' in a deterministic world and deliberation is generally effective, not futile (p.106).

(3) The sort of fatalistic circumstances that the refutation argument is concerned with (i.e. situations that concern the 'causal impotence' or 'futility' of deliberation - *Elbow Room*, 15,104,106) may nevertheless vary in significant ways. Compare, for instance, Dennett's 'bogeymen' examples such as being controlled by 'the Peremptory Puppeteer' and 'the Hideous Hypnotist' (*Elbow Room*, 8-9). As Dennett points out, the phenomenology of agency/fatalism is very different in these cases.

(4) There are exceptions to this generalization. See, e.g., Isaiah Berlin, *Four Essays on Liberty* (Oxford: Oxford University Press, 1969), xiii. Although Berlin accepts the refutation argument and its associated understanding of fatalism, he nevertheless argues for the responsibility-incompatibilist claim on independent grounds.

(5) Some incompatibilists, of course, object to deterministic metaphysics on the ground that it implies 'mechanism', and this is incompatible with the sort of purposive explanations that are essential to responsible agency. This distinct and more radical line of incompatibilist reasoning (which Dennett labels as worries about 'sphexishness'; *Elbow Room*,10-14) is not essential to their position. On this see Watson's introduction to Free Will, 11-14.

(6) For a discussion and interpretation of the relevance of the origination/contribution distinction for the free will debate see Robert Nozick, *Philosophical Explanations* (Oxford: Oxford University Press, 1981), 313. Nozick interprets fatalism as denying that our actions have any 'contributory value', and the problem of causal determinism as the suggestion that our actions would be left without 'originatory value'.

(7) For an influential and illuminating discussion that articulates these incompatibilist intuitions see Thomas Nagel, 'Moral Luck', reprinted in Gary Watson, ed., *Free Will* (Oxford: Oxford University Press, 1982); esp. 183 on 'genuine agency' and 'shrinking' responsibility. Another similarly important and interesting discussion of these matters is presented in Gary Watson, 'Responsibility and the Limits of Evil: Variations on a Strawsonian Theme', reprinted in J. M. Fischer & M. Ravizza, eds., *Perspective on Moral Responsibility* (Ithaca & London: Cornell University Press, 1993); esp. 143-44 on 'origination' and the 'historical dimension' of responsibility. Both Nagel and Watson (consistent with usual incompatibilist concerns) emphasize the relevance of worries about 'origination' for issues of responsibility.

(8) It is no coincidence, for example, that Dennett's account of responsibility is wholly pragmatic and forward-looking in character (*Elbow Room*, 156-65). On this see Gary Watson's review of *Elbow Room* in the *Journal of Philosophy*, 83 (1986), 517-22.
(9) Susan Wolf, *Freedom Within Reason* (New York & Oxford: Oxford University Press, 1990), 40-45.
(10) See, in particular, Nagel's remarks on 'the contributions of fate' and their tendency 'to erode most of the moral assessments we find it natural to make' ('Moral Luck', esp.176,180,182). I note in passing that not all incompatibilists would accept that their position should be interpreted in terms of concerns about 'origination'. Some, for example, may articulate their incompatibilism in terms of the issue of 'alternate possibilities' or 'freedom to do otherwise'. Incompatibilist concerns of this nature, however, depend on a particular ('categorical') interpretation of these requirements which on analysis, it may be argued, reflect (deeper) concerns about origination. It suffices, in any case, that concerns about origination constitute a standard incompatibilist perspective on the free will issue. For the purpose of concise presentation, therefore, I will not elaborate on these complexities.
(11) P.F. Strawson, 'Freedom and Resentment', reprinted in Watson, ed., *Free Will*, 59-80. Strawson's paper is also reprinted in Fischer & Ravizza, eds., *Perspectives on Moral Responsibility*; see also the editors' introduction for a helpful discussion of various responses and criticisms of Strawson (pp.4-25).
(12) I develops this line of criticism of Strawson in 'Strawson's Way of Naturalizing Responsibility', *Ethics*, 102 (1992), 287-302 (see esp. pp.296-97,300-01). See also Watson's related discussion of Strawson's difficulties in accounting for 'exempting conditions'; 'Responsibility and the Limits of Evil', esp. 125-26.
(13) Dennett's *Elbow Room* is an important contribution to this aspect of contemporary compatibilist thinking. Other influential accounts of this kind include Harry Frankfurt, 'Freedom of the Will and the Concept of a Person', and Gary Watson, 'Free Agency', both reprinted in Watson, ed., *Free Agency*. Closely related to the second strand of naturalized responsibility is the issue of 'reflexivity' and 'reason-responsiveness'. Dennett, among others, devotes considerable attention to this matter. See esp. *Elbow Room*, Chp.2.
(14) There is considerable variation in the specific ways that these two strands of naturalized-responsibility are developed and articulated. Compare, for example, the Humean way of developing and blending these themes as presented in Paul Russell, *Freedom and Moral Sentiment: Hume's Way of Naturalizing Responsibility* (New York & Oxford: Oxford University Press, 1995), with the essentially Kantian account presented in R. Jay Wallace, *Responsibility and the Moral Sentiments* (Cambridge, Mass.: Harvard University Press, 1994). It is also important to note that not all contemporary compatibilists accept both of the two strands described above (see, e.g., note 8 above on Dennett).
(15) Some incompatibilists may argue that their pessimism about the fatalistic implications of determinism are not entirely based on worries about responsibility, although this is their primary concern. In so far as incompatibilists have fatalistic concerns independent of the issue of responsibility they share common cause with compatibilist-fatalists - as I will explain.
(16) This terminology is a prominent feature of Strawson's discussion in 'Freedom and Resentment', where it is used to describe the positions of the major parties in the free will dispute: incompatibilists being 'pessimists', compatibilists being 'optimists'.
(17) In Elbow *Room Dennett* interprets his own defence of compatibilism as a vindication of 'optimism' over 'pessimism' (Elbow Room,18-19,169). His discussion makes clear that from an orthodox-compatibilist perspective incompatibilist claims are not innocuous, as they generate

negative emotions such as 'fear', 'anxiety', 'dread' and so on. Dennett's general conclusion is that all such incompatibilist 'pessimism' can be effectively discredited and shown to be motivated by various kinds of (philosophical) confusion and/or illusion. This includes, notably, pessimism about fate.

(18) Spinoza, *Ethics* (Indianapolis & Cambridge: Hackett, 1992),102 (Part III, Preface); translation by Samuel Shirley.

(19) The metaphor of being 'governed by nature' may be taken to suggest that Nature (somehow) 'controls' us for its own ends and purposes. This would involve confusion and should be avoided. For this reason it is important to distinguish worries about origination-fatalism from worries about supernatural-fate. In the case of supernatural-fate it is argued, not only that the ultimate source of character and conduct does not lie with the agent (and thus has an external source) but, moreover, the external source is some supernatural agent or cosmic being who 'manipulates' or 'directs' our (human) lives according to some (alien) design or plan. Worries about loss of 'sovereignty', however, need not presuppose any such 'bogeyman' to be at work. In general, there is no reason to suppose that a mistake of this kind is required to motivate pessimistic concerns about origination-fatalism. (One of the unsatisfactory aspects of Dennett's efforts to defuse worries about fatalism is that he tends to assimilate worries about origination with worries about supernatural-fate: see *Elbow Room*, 7-17, and Chp.3.)

(20) The expression is Dennett's (*Elbow Room*, chp.4, esp. p.100) and it is indicative of the extent of his 'optimism' on such matters. See also the papers by Frankfurt and Watson cited in note 13 above.

(21) Dennett notes himself (*Elbow Room*, 85,156) that 'a completely self-made self, one hundred per cent responsible for its own character, [is] an impossibility'. The question arises, however, what percentage is required for a 'self' to be 'self-made' - will any percentage do? It should also be noted that Dennett does not claim that we avoid worries about fatalism to the extent that we are 'self-made-selves'. On the contrary, since he accepts the (classical) refutation argument, and its narrow conception of fate as contributory-fate, all that is required to avoid worries about 'fate', he claims, is for deliberation and action to be causally effective.

(22) There are variations on this general problem in compatibilist literature. Wallace, for example, suggests that 'powers of reflective self-control' constitute the relevant moral capacities required for responsible agency. (See the discussion concerning moral capacities above.) These powers, he says, involve the possession of the ability to grasp and apply moral reasons, and to regulate behaviour on this basis. (*Responsibility and the Moral Sentiments*, 157) However, as Wallace concedes, agents may possess these powers and yet have no ability to determine the way that they are exercised in particular circumstances (pp. 180-94, 201-14). This is, however, precisely what is required for 'sovereignty'. Hence, even if Wallace's defence of responsibility-compatibilism is accepted, the concerns about origination-fatalism remain unanswered.

(23) For a brief account of this matter see Russell, *Freedom & Moral Sentiment*, 128-30.

(24) In respect of this, consider Watson's illuminating and suggestive reflections on the significance of the case of Robert Harris. ['Responsibility and the Limits of Evil', 137-46]. Harris was a notably brutal Californian killer (i.e. when viewed as a 'victimizer') and also the product of an exceptionally brutal childhood (i.e. when viewed as a 'victim'). Watson interprets the significance of the 'historical' considerations relating to Harris's childhood and moral development in terms of their tendency to influence our reactive attitudes (i.e. to produce 'ambivalence': 137-38). There is, however, another way of looking at this case, more in keeping with compatibilist-fatalism. Reflection on such circumstances press the thought upon us that who we are, and what we are

responsible for to other human beings, depends ultimately on factors that we have no control over. These reflections are even more troubling when, as Watson puts it, we 'turn our gaze inward' and recognize 'that one's moral self is such a fragile thing' (p.139). In contrast with this view, orthodox-compatibilism suggests that historical considerations of this kind are untroubling so long as they do not discredit or dislodge our (natural) commitment to reactive attitudes.

(25) It is, in particular, a notorious stumbling-block of libertarian metaphysics that it is unable to make clear what is required for 'genuine agency' beyond simple indeterminism. The difficulties facing the libertarian are well described in the closing passages of Nagel's 'Moral Luck'; Nagel's *The View from Nowhere* (New York & Oxford: Oxford University Press, 1986), Chp.7; and also Galen Strawson, *Freedom and Belief* (Oxford: Clarendon Press, 1986), Chp.2.

(26) The many difficulties associated with making sense of the thesis of immortality are well-known. An interesting discussion of the desirability of immortality is presented in Bernard Williams, 'The Makropulos Case', in *Problems of the Self* (Cambridge: Cambridge University Press, 1973).

(27) 'We cannot look squarely at either death or the sun.' LaRochfoucauld, *Maxims*, No.26.

(28) The only way to evade these pessimistic reflections about origination-fatalism is to provide some (coherent) account of 'genuine agency' that is premised on indeterministic metaphysics. For recent libertarian efforts along these lines see the various papers in Timothy O'Connor, ed., *Agents, Causes & Events: Essays on Indeterminism and Free Will* (Oxford: Oxford University Press, 1995).

(29) Consider, for example, Watson's discussion of Robert Harris, as cited note 24 above.

(30) The compatibilist-fatalist, as explained, interprets the specific way that responsibility and fate "come fused together in human life" in terms of (rejecting) the exclusion thesis, and distinguishing between origination and contributory fatalism.

(31) Dennett associates the pessimism generated by the 'bugbear' of fatalism with the condition of 'puppets' or 'dolls' - something that really is a 'terrible condition' (*Elbow Room*, Chp. 1).

(32) John Macquarrie, *Existentialism* (Harmondsworth, Middx., Penguin, 1973), 191.

(33) This sense of the contingency of human existence, and its relevance to our view of ourselves as (responsible) agents who are nevertheless 'thrown' into our own particular circumstances, is something that many moral theories (most notably Kantianism) strongly resist. On this see Bernard Williams, 'Moral Luck: a postscript', reprinted in *Making sense of humanity* (Cambridge: Cambridge University Press, 1995), 246.

(34) My comments here draw on Thomas Nagel's influential discussion of the our sense of the absurd as it relates to human life: 'The Absurd', reprinted in *Mortal Questions* (Cambridge: Cambridge University Press, 1979).

(35) Despite this, some may be tempted to question the freshness of compatibilist-fatalism on the ground that each of its two component claims are (very) familiar. It should be clear, however, that the particular interest of this position does not rest with its two component claims considered in isolation from each other, but rather with the effort to combine two claims that have traditionally been treated by both the major parties in the free will dispute as incompatible - a thesis which compatibilist-fatalism rejects. I am unaware of any compatibilist thinker who has defended the 'mixed' optimist/pessimist position of compatibilist-fatalism as described. See, however, Saul Smilansky, 'Does the Free Will Problem Rest on a Mistake', *Philosophical Papers*, 22 (1993), 173-88. Smilansky pursues themes that are very relevant to the position taken in this paper.

(36) The qualifying clause in this sentence (i.e. 'or accept the legitimacy of') provides scope for the weaker position that allows that some compatibilists, after due reflection, may remain untroubled

by any considerations regarding origination. (See section V above.) On the assumption that there is no confusion about the source and quality of the pessimism at issue, nor any failure of due reflection in such cases, but only a divergence of sensibility, then orthodox-compatibilism may be judged no less - and no more - legitimate than compatibilist-fatalism. As I have indicated, however, it may be argued that a failure to be troubled by considerations regarding origination is best explained in terms of a lack of appropriate reflection, and that a suitable sensibility can be cultivated on the basis of such reflection.

(37) Incompatibilists, of course, remain committed to the exclusion thesis in so far as it is essential to their defence of the responsibility-incompatibilist claim.

An Abstract of

COMPATIBILIST-FATALISM
by
PAUL RUSSELL, University of British Columbia

International Conference on
Moral Responsibility and Ontology
Utrecht, Netherlands: 4-7 June 1998

In this paper I argue that compatibilists must embrace a richer conception of fatalistic concern; one that recognizes the legitimacy of incompatibilist concerns about the origination of character and conduct. Incompatibilists articulate these concerns in terms of worries about ("deep") responsibility -- something that compatibilists clearly reject. Nevertheless, as I explain, pessimistic concerns about ("genuine") agency survive the (assumed) success of compatibilism with respect to responsibility. I consider a variety of efforts to refute or discredit pessimistic concerns of this nature and show that none of them are successful. On this basis, I conclude that any plausible compatibilist position must allow that determinism has fatalistic implications of some significant and relevant kind, and accept that agents may be legitimately held responsible in circumstances where they are subject to fate. The position generated by these compatibilist concessions to incompatibilism is "compatibilist-fatalism".