



Compatibilist Libertarianism: Why It Talks Past the Traditional Free Will Problem and Determinism Is Still a Worry

ABSTRACT: *Compatibilist libertarianism claims that alternate possibilities for action at the agential level are consistent with determinism at the physical level. Unlike traditional compatibilism about alternate possibilities, involving conditional or dispositional accounts of the ability to act, compatibilist libertarianism offers us unqualified modalities at the agential level, consistent with physical determinism, a potentially big advance. However, I argue that the account runs up against two problems. Firstly, the way in which the agential modalities are generated talks past the worries of the incompatibilist in the traditional free will problem. As such, it fails to dispel the worries that determinism generates for the incompatibilist. Secondly, in spite of the ingenious use of the supervenience thesis and multiple realizability, the position still allows us to generate the old worry that determinism at the physical level would mean no alternate possibilities at the level of agency. In particular, I develop a new example, the ‘atomic slit case’ that demonstrates how physical level information is salient to what is possible at the agential level, motivating incompatibilism.*

KEYWORDS: alternate possibilities, compatibilism, moral responsibility, free will, compatibilist libertarianism, supervenience, Christian List

Introduction

The question of whether alternate possibilities for action are compatible with physical determinism invites two responses from the compatibilist. Firstly, they can try to show that the ability to do otherwise, suitably understood, is consistent with physical determinism. Secondly, they can argue that if alternative possibilities are not required for morally responsible action, then physical determinism is no threat. The first route, sometimes called *leeway* compatibilism, comes in its classical form with David Hume’s discussion in the 1777 *Enquiries Concerning Human Understanding and Concerning the Principles of Moral* (Hume 1993: 53–69). In the twentieth century, this approach was further developed in the work of

I would like to thank Maria Alvarez, Tom Pink, Christian List, Fearghus Horan, Alex Geddes and Nadine Elzein for helpful discussion and critical feedback on earlier drafts of this paper. Thanks also to the editorial team at *The Journal of the American Philosophical Association* and Ellen Wert, the Cambridge University Press copyeditor who processed the manuscript.

G. E. Moore (1903), A. J. Ayer (1954), J. J. C. Smart (1961), later by David Lewis (1981) and Bernard Berofsky (2005, 2012) and is sometimes called ‘conditional’ or ‘traditional’ compatibilism. More recently in the leeway camp, a sophisticated reworking of the old conditional approach is the so-called new dispositionalism of Michael Smith (1997), Michael Fara (2008), and Kadri Vihvelin (2004, 2013).

In contrast, the second approach, sometimes called *source* or *causal history* compatibilism, was developed in the wake of the ‘Frankfurt example’ strategy, after Harry Frankfurt (1969) argued alternate possibilities are not necessary for morally responsible action. Where does compatibilist libertarianism—the term coined by Christian List (2014, 2019)—fit into this picture? The theory provides a way forward for those rejecting the causal history approach—that is, those who wish to retain a central place for alternate possibilities, but who at the same time reject the existing leeway accounts. This is a well-motivated direction of travel, as both the old conditional compatibilism and the new dispositionalism have been criticized for being at odds with ordinary usage by Roderick Chisholm (1964), Keith Lehrer (1968), Randolph Clarke (2009), and List (2014, 2019). The causal history approach is likewise controversial. If successful then, we would gain a compatibilism that does not derive its naturalistic respectability by distorting the meanings of ‘could’ in ordinary usage. The general approach has some precursors, and something like it is taken up by Christopher Taylor and Daniel Dennett (2005), but it has recently been systematically developed by List (2014, 2019). In what follows, I deal exclusively with List’s account as it provides the most detailed version of this approach.

List argues for his position in a twofold way. Firstly, his ‘bottom-up’ argument provides a model of how agential possibility and physical possibility are related, explaining how indeterminism at the agential level is compatible with determinism at the physical level. Secondly, in his ‘top-down’ argument, he appeals to an ontological principle from the philosophy of science, the naturalist ontological attitude, to argue that agential possibility as he outlines, is both real and the salient notion we should be using in discussions of free will because our best theories of human agency require the notion. I find List’s account problematic in at least two respects: (1) the position does not sufficiently address the issue that free will understood as a type of control, must be sufficient for attributions of moral responsibility to be justified; and (2) the account appears to have difficulty dealing with our ordinary commitments concerning how we conceive of ourselves as agents related to our physical bodies, commitments that remain in tension with physical determinism.

1. List’s Account

List combines a number of claims which together enable him to argue for the position that physical determinism would not deprive us of the variety of alternate possibility we need for free will and responsibility. Before examining those, I revisit a simple argument for incompatibilism in order to set the context for List’s point of departure. List gives the basic form of argument for leeway incompatibilism as follows:

Premise 1. A necessary condition for someone's action to count as free is that the agent can do otherwise.

Premise 2: Determinism implies that the agent cannot do otherwise.

Conclusion: Either there are no free actions, or determinism is false (or both). (List 2014: 156).

The same reasoning underpins Peter Van Inwagen's (1983: 70) consequence argument. List does not challenge the first premise. He argues that premise two is false even when we interpret the ability to do otherwise in an unqualified modal way and not conditionally or dispositionally. With this in mind, List specifies the basic argument more precisely:

Premise 1: Free will requires that (at the time of interest) more than one alternative course of action is possible for the agent.

Premise 2: Determinism implies that (at the time of interest) only one alternative course of action is possible *for the agent*.

Conclusion: Free will and determinism are incompatible. (List 2014: 160, my emphases.)

List says that the thesis of determinism is a claim about *physical* possibility. The truth of physical determinism does not entail the truth of the second premise here as that premise is about what is possible for an *agent*. Premise two above would be true if a certain linking assumption was true connecting physical and agential possibility. It would have to be the case that if at a given time only one sequence of events was physically possible, then, as a consequence of that fact, only one course of events would be possible for the agent. List says that this is not generally true. There is a conceptual difference between these two types of modality, agential and physical, they cannot be assumed to run together. Incompatibilists must provide reasons for why they have to run together. List gives us reasons why they do not.

List provides two arguments that, taken together, support the conclusion that agential alternate possibility is the salient modality in play and that the agential modalities are not threatened by physical determinism: firstly what he calls the 'bottom-up' argument and secondly what he calls the 'top-down' argument.

The bottom-up argument is a toy model that explains how the two levels of description (microphysical on the one hand, agential on the other) are connected. (See List 2014: 162–67 for the full statement of the toy model.) The main feature of this first argument is the thesis of multiple realizability about mental states (that is, agential states) from physical states. List summarizes:

[E]ach agential state corresponds to an equivalence class of physical states, consisting of all those physical states that give rise to that same agential state. It is precisely this coarse-grained nature of someone's agential state that can render it consistent with multiple distinct

physical sequences of events, each of which “passes through” one of the possible physical states realizing the given agential state.

As long as some of these sequences correspond to different courses of action, it follows that more than one course of action is *possible for the agent*, contrary to Premise 2. (List 2014: 162)

List presents diagrams (figures 1 and 2) to illustrate how multiple realizability makes possible branching (alternative possibilities) at the agential level while determinism with no branching is the case at the physical level.

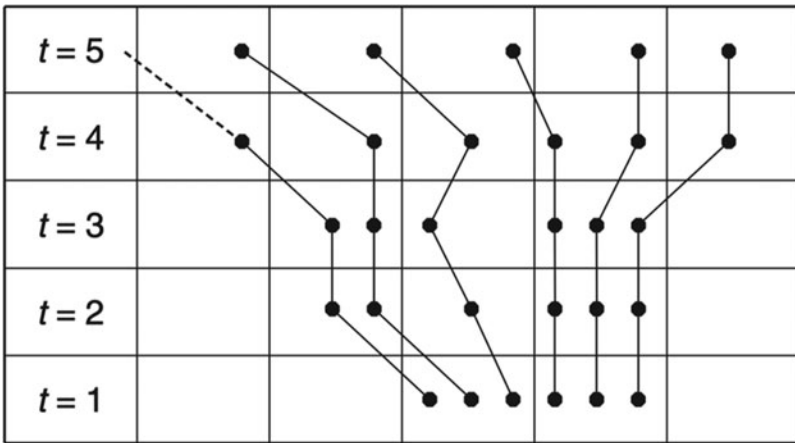


Figure 1. —World histories at the physical level. Source: Christian List. (2014) ‘Free Will, Determinism, and the Possibility of Doing Otherwise’. *Noûs*, 48, 156–78, at 166. Used by permission.

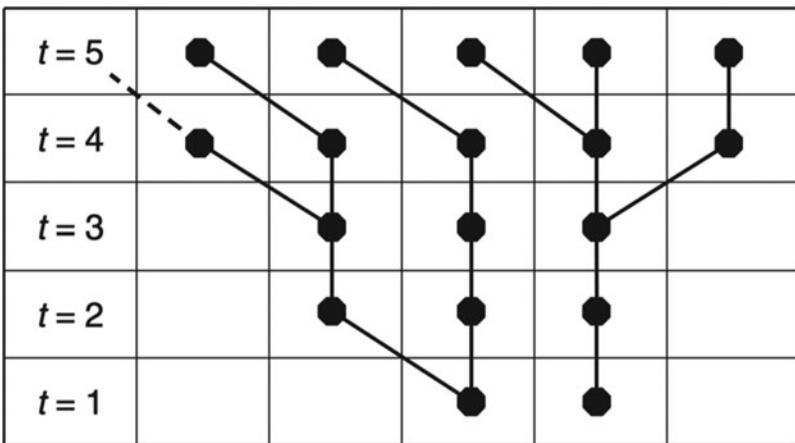


Figure 2. —World histories at the agential level. Source: Christian List. (2014) ‘Free Will, Determinism, and the Possibility of Doing Otherwise’. *Noûs*, 48, 156–78, at 166. Used by permission.

On the left hand side of [figure 2](#), looking at the agential state at time t_3 , the agent has a choice between two different courses of action at t_4 . However, the corresponding physical supervenience bases that underwrite this state of affairs are non-branching, as can be seen from the lines passing through t_3 and t_4 in [figure 1](#). I agree that the toy model succeeds in showing how, in principle, there can be a system whereby an emergent higher level is indeterministic and yet supervenes on a lower level that is deterministic. However, the incompatibilist can now ask whether the modal semantics of agential possibility List describes is coextensive with the modal semantics of our ordinary practice and language. In other words, when we ordinarily make claims about what agents can and cannot do, are the truth conditions of those claims the same as the truth conditions for the agential level as List describes it in the toy model? Specifically, are the truth conditions of those claims appropriately isolated from physical determinism? This is where List's second argument comes in. In the top-down argument, List claims that physical possibility is an inappropriate level of description, or frame of reference, when we are talking about what is possible for an agent:

When we are interested in whether a particular action is possible for an agent, by contrast, the appropriate frame of reference is not the one given by fundamental physics, but rather the one given by our best theory of human agency. Thus the description of the world that matters here is not a (microscopic) physical one, but a (macroscopic) psychological one. Candidate theories that provide the right level of description include some advanced versions of psychological decision theory, such as those we find in economic psychology or cognitive science, which are currently our best attempts to make scientific sense of intentional agency. In fact, even folk psychology outperforms physics or neuroscience when it comes to understanding and explaining human behaviour across different domains and outside isolated laboratory conditions. (List [2014](#): 161–62)

A generally accepted method for establishing whether certain entities or properties are real is to consult our best scientific theories of the relevant domain. If postulating the entities or properties is indispensable from the perspective of those theories—that is, the theories couldn't explain their target phenomena without postulating those entities or properties—then this gives us good reasons to accept the entities or properties as real. (List, Caruso, and Clark [2020](#): 63)

The core move here is the appeal to the natural ontological attitude. (See Quine [1977](#); Fine [1984](#): 83–107.) The natural ontological attitude is the idea that we should posit the necessary commitments of our most successful theories as real. For example, in fundamental physics, if electrons and quarks are a necessary part of the scientific model that allows us to most accurately explain what is going on in the physical world then electrons and quarks really exist. Similarly, when we are modeling human behavior, the most successful theories are decision or social choice theory.

Crucially, those theories posit the ability to select options from sets of possibilities on the basis of rational deliberation. That feature is a necessary part of the theoretical apparatus that allows us to best model and make sense of human behavior. Analogously then, the reality of List's 'agential possibility' is brought about. Free will is a higher-level phenomenon, with an ontology akin to that of desire and belief and no less real than those. List says, "Free will, in the technical sense of an agent's having a choice between more than one course of action in many situations, is a key presupposition of our best scientific theories of agency, at least when these theories are understood literally" (List 2014: 168). In summary, it is *because* the agentive modalities are a necessary part of our most successful theories of human behavior that we should look to those modalities when asking questions about what agents can and can't do and not look to (microscopic) physical possibility. The microphysical scientific reference frame is not one where we have anything like success in modeling human behavior. It appears virtually impossible to explain even basic behavior in microphysical terms. You get lost in information overload and miss the key difference making phenomena. The content of microphysical statements does not translate into propositions in terms of intentional agency and it looks like there can be no bridge laws between levels either. All of this is what underwrites List's claim that we *should* work with agentive, not microphysical modality.

List's case consists of these two arguments, bottom-up and top-down, taken together. The top-down argument is supposed to establish that the notion of agential possibility is the appropriate one to work with when we're interested in settling questions about what agents can and cannot do. The bottom-up argument provides a formal framework wherein List makes clear how the indeterminism of the higher agential level can supervene on the determinism of the lower microphysical level.

In response, I agree List has successfully shown how in principle indeterminism can exist at a higher-level frame of reference that nevertheless supervenes on a lower-level frame of reference that is deterministic. However, once again, the question is whether the semantics of that higher-level agential indeterminism List argues for is what we need to speak to the traditional free problem.

In light of this question, I focus on two issues. Firstly, I examine the status of the free will List can generate given our best theories of human behavior. I argue that it does not follow that *this* kind of free will is appropriately connected up with the conditions on moral responsibility, as any concept of free will must be if it is to be relevant to the core issue at the heart of the traditional problem of free will. There is a gap in the top-down argument that must be plugged. Even granting the reality of free will generated by the top-down argument, we can sensibly question whether that sense is sufficient to constitute the control condition on moral responsibility. This worry persists when we consider folk psychology as a working theory of human behavior in addition to the more formal decision theory.

This first argument that List's agential possibility is not appropriately connected up with the conditions on moral responsibility is similar to criticisms made by Nadine Elzein and Tuomas Pernu (2017). I developed this argument independently in Wright (2016).

Secondly, I provide a new argument to show that the agentive modalities as List defines them are not coextensive with how we standardly talk about possibility at the agential level. Specifically, this *physical location argument* shows how facts at the microphysical level of description are pertinent and constraining on what can be said to be possible for us as agents. If this argument is sound, then the incompatibilist will be in a position not only to claim that List's 'agentive possibility' might not be sufficient for attributions of moral responsibility, but in addition, that our everyday concept of the ability to do otherwise is still incompatible with physical determinism.

2. The Problem with List's Top-Down Argument

List's top-down argument does not provide us with moral responsibility. Therefore, it talks past the traditional free will problem. The ontology of free will delivered by adopting the natural ontological attitude toward social choice or decision theory (in the top-down argument), will be no thicker than the theories making use of the concept require, such that they constitute successful theories. Given that social choice theory is trying to model how and why we act given our desires and goals, it seems uncontentious that the modal notions it makes use of need not make reference to microphysical possibility. The coarse-grained descriptions of the agential level are sufficient and microphysical descriptions are in fact inappropriate for this kind of modeling, not merely unnecessary. The relevant agential level concepts such as 'choice over option sets' can enable the theories in question to work to their full explanatory potential without needing to be anything metaphysically over and above a higher-level agential-psychological phenomenon. If we are trying to model the rationality of behavior, we can do that successfully with a thinner agential (and compatibilist) notion of choice over alternatives. List is surely right about this.

The worry is that the special science theories that posit free will (in List's sense of agential possibility over alternatives), do not try to model our practice of holding agents morally responsible. That is to say our blaming practices and framework of reactive attitude ascriptions. Hence, more argument is needed to establish that List's ontology of freedom should stop people worrying about the threat determinism might pose to moral responsibility. As noted above, what social choice theory does do is model the rationality of agents given certain goals and desires they have. But why should we expect that the theoretic entities such special science theory commits us to would also be related in the right sort of way to our notion of moral responsibility as it is embedded in the web of our standing ordinary commitments and practice? The problem is that those models (that is, social choice theory), do not seem to make any assumptions one way or another regarding whether the sense of choice at issue in them is also the kind of choice that would be sufficient for the control condition on moral responsibility. In summary, it may well be true that decision theory uses a notion of choice that commits us to the reality of alternate possibilities. In that sense, List may well be right that we can be said to be free, genuinely to have free will in that sense. All of this said, it does not follow that this notion of free will is appropriately linked

with moral responsibility. Why? The incompatibilist is entitled to reply that there are conceptual commitments as part of our folk theory and ordinary language that are nevertheless incompatible with determinism. Whether or not this is in fact true, List's argument would not be affected. But it is precisely on such a point that the traditional free will problem hangs. In other words, List is talking past the crux of the traditional problem.

List might reply that we *should* be using his agential possibility when talking about these matters. In reply, the incompatibilist is entitled to say that this depends on our aims. On the one hand, when modeling behavior and explaining outcomes, List is right to say that we need the agential sense he defines and should indeed be using it in order to model and make sense of what we see going on at the higher levels of reality. If we do not, we will miss the key difference making phenomena and patterns. On the other hand, if we are wondering whether we have the kind of control sufficient for ascriptions of moral responsibility, then that is a different project and may need something in addition to List's agential possibility. The success of the explanatory and modeling project in the special sciences does not require any deeper metaphysical commitments. But the traditional worry at the heart of the free will problem *just is* that such commitments might be tacitly required as a condition on moral responsibility. In other words, the normativity behind List's claim that we should be using the agential notion does not engage with the incompatibilist worry.

It might be said in response to my worry above, that List's ontology of free will is not merely generated from the technical sense of alternative possibility utilized by social choice theory. As List himself emphasizes, folk psychology itself is a very successful theory that models human behavior and it, too, posits alternate possibilities. If that is the case, given we are adopting the natural ontological attitude, do not we immediately arrive at an ontology for our folk psychological sense of alternative possibility *as it is connected up with responsibility and blame*, hence 'plugging the gap' I claimed was present when we were only concerned with the more scientific decision or social choice theory? The thought is that since it is part of our folk picture of human behavior that people are blameworthy only when they can do otherwise, if we apply the top-down argument to folk psychology then we get the ontology of our folk conception of free choice delivered up and *that* concept is appropriately connected with moral responsibility.

On reflection however, it appears that the ontology of free will generated with respect to folk psychology is subject to a similar worry. When we treat folk psychology as a working theory that models human behavior, having adopted the natural ontological attitude, just like before with the more technical social choice theory, we only get an ontology of folk free will sufficient to *explain* human behavior and practice but that is not necessarily the same thing as getting an ontology of free will sufficient to *justify* aspects of human behavior (namely our moral responsibility practices and reactive attitudes). Consider the example of religion. If we are trying to construct an anthropological theory of belief to explain some pattern of ancient religious behavior, the successful model can allow us to do just that without having to be in any way a justification of those religious practices.

The problematic gap here is still the potential difference between the ontology of free will that is delivered up by any successful theory *of* human behavior on the one hand, and the semantics and inferential role of ability claims and moral responsibility attributions *in* ordinary practice on the other. This is because not everything we might be metaphysically committed to in our ordinary practice concerning freedom and responsibility needs to be realizable or exist in order for even a folk theory to model our behavior and practice successfully. Nevertheless, such commitments may still be necessary for moral responsibility ascriptions to be justified. If such commitments are in fact indeterministic, we remain in the grip of the traditional incompatibilist worry.

It is sometimes hard to see this point because when we are talking about alternative possibilities in ordinary discourse or folk psychology, these surely *just are* the possibilities that List means right? And surely these possibilities are just what we are talking about when we talk about (one of) the control conditions on responsibility. So what is my point? Once again, it is possible that List is elevating that part of the semantics of everyday ‘could’ and ‘can’ talk that has the semantics he defines for the reasons he gives—that is, because it is necessary for our models of behavior. All well and good. There may however be additional content that is not elevated, but which is a necessary condition on moral responsibility. Nothing List says rules out this possibility.

It is helpful to consider modeling human behavior from an outside, alien perspective. From their observations of us and the modeling they might undertake, whether it makes use of more technical social choice theory or is constructed out of our existing folk terminology, we get an ontology only for the theoretic entities at issue sufficient to explain what the outside party needs in order to understand our behavior when they look at us going about our business on Earth. That ontology might even require a detailed incompatibilist commitment on our part in the form of our having various beliefs about control, alternative possibilities and their supposed connection with moral responsibility ascriptions. But even if all that incompatibilist conceptualization was in fact necessary for the outside perspective to adequately model us, an actual ontology for the supposed *objects* of those incompatibilist beliefs would not be required and therefore it still would not follow that any of our moral responsibility ascriptions were in fact justified by their own lights. Why? Because explaining behavior is not the same thing as justifying it. If, on the other hand, we want a theory that would tell us what ontology is sufficient to justify moral responsibility ascriptions, then we are simply back in the traditional free will debate itself—that is, the debate about what kind of metaphysical power and control we need to possess in order to be responsible.

I have tried to show it does not follow from the fact that we can derive an ontology of free will in the way List describes, that that sense of free will is coextensive with our ordinary or folk notion of free will as it is connected up with the concept of moral responsibility. That said, List does very briefly discuss the ordinary language issue. He notes that the present analysis is broadly consistent with some accounts of the ordinary meaning of ‘can.’ He goes through some examples from Angelika Kratzer (1977) and mentions work by John Maier (2015) on the agentive modalities. In response, it is important to remember that the Kratzer semantics

itself does not explicitly settle the argument between compatibilist and incompatibilist (see Weir 2016 on this point). Maier (2015: 133) also makes the point that his analysis does not directly speak to the question about compatibility with determinism, though it is an important framework within which the metaphysical debate can be set up. So it remains to be seen whether the sense of agential possibility captured by List is the sense of possibility we are after. Perhaps it might be said that because List has put a compatibilist modal semantics on the table, the ball is in the incompatibilist court. I concede that would indeed be the case if List's modal semantics had been derived from an analysis of ordinary language and practice but it is not. Furthermore, as I demonstrate below, we still have good reason to think that the agential modal semantics List develops is not coextensive with our ordinary usage in a problematic way that points towards incompatibilism.

It is worth noting the distinction between two types of compatibilist approach: diagnostic and prescriptive. It seems that List is not clear which of these camps (or both) his project falls into. On the one hand, the methodology of using the natural ontological attitude might be seen as falling under the prescriptive heading. List is telling us we have reason to adopt a certain sense of free will independent of the analysis of ordinary language. On the other hand, he does engage in diagnostic work here, precisely because he discusses the Kratzer semantics and is at pains to claim this is evidence that his analysis accords with ordinary usage. Toward the end of the essay, List discusses what we should do if it was discovered that determinism were true and we woke up to the morning papers with headlines announcing the fact. In that circumstance, after considering options like ceasing with responsibility practices and the institution of punishment or in fact carrying on much as before, he says:

Surely we would do the latter: giving up our conventional understanding of free will and revising the very fabric of how human society works would be an overreaction. The approach to free will offered in this paper shows why this is so. The mildest revision of our technical vocabulary—namely the shift from physical to agential possibility in the analysis of free will—is sufficient to rehabilitate practically everything we conventionally believe and say about free will, even against the background of determinism. For this reason, my proposal seems to have common sense on its side.

In conclusion, I suggest that the best way to defend the compatibility of free will and determinism is to recognize that free will is not a physical phenomenon, but a higher-level phenomenon on a par with other familiar higher-level phenomena such as beliefs, desires, and intentions. If we are searching for free will at the level of fundamental physics, we are simply searching in the wrong place. (List 2014: 174)

List's phrase 'mildest revision to our technical vocabulary' is telling here. Revision would imply that List's concept of free will is not exactly the same as the one we

have been using. If he is happy to accept that the concept he advocates is not the same as we find in ordinary practice, why is it so important to discuss the Kratzer semantics? Still, presumably the milder the revision the better here and so just because a project is revisionist that does not mean that concern with ordinary usage evaporates altogether. However, even if such revision is mild, we still need to know whether the revision licenses the continued deployment of the family of moral concepts and practices that free will is taken to be connected up with.

The central issue in the traditional free will problem is whether in a deterministic universe we could have the type of control necessary to justify ascriptions of moral responsibility. Traditional compatibilist analyses of ‘could have done otherwise’ in ordinary language automatically resolve this worry about responsibility (if they are successful). This is because the ordinary language concept of moral responsibility is linked to our ability to do otherwise in such a fashion that if we can be shown to have the latter (in a deterministic world) then we may (under the right conditions) be justified in ascribing responsibility to particular actions and persons. In ordinary language analysis, if we have shown something about free will then we have also, at the same time, shown something about a necessary condition of moral responsibility. But, in close analogy with what I have said above, the free will List has argued for must still be shown to be sufficient for moral responsibility because it is not obvious that it is the same as that necessary condition of responsibility found in ordinary discourse.

This worry remains even for those who are not concerned about whether List’s free will is coextensive with our ordinary language notion. Why? Because you still have to show that a concept of free will is sufficient to hold someone responsible even if you are a revisionist about the concept of responsibility. A person sympathetic to revisionism about free will and responsibility indeed must still question whether it makes sense to hold people morally responsible for their actions in a physically deterministic world.

I am not saying that ordinary language analysis is crucial for any successful argument in this domain. Another way of proceeding would be to establish that the free will generated by List is connected up in the right kind of way with our practice of moral responsibility as we *should* conceive it (even if that too is a revised concept). However, I take it that List does not see himself as undertaking this kind of revisionist project as a response to the traditional problem. That said, we might concede that moral practice as it stands requires the truth of indeterminism but that we can and should embrace a revised responsibility system that is compatibilist, making use of List’s agential possibility.

In personal correspondence, List has responded to my criticism that his free will is not appropriately connected with the more widely held notion of moral responsibility. He contends that notions of agency and responsibility belong to the agential level, not the fundamental physical one. It is intentional-agency discourse that feeds into debates about responsibility, not particle-physics talk. As a result, List argues, it is appropriate to assess the empirical premises for various claims about responsibility at the agential level, not the fundamental physical one. In response, I claim this does not speak to the point. I agree that it is intentional agency level talk that feeds into debates about responsibility, specifically the

requirement for alternative possibilities as a condition on morally responsible action. But, as I have tried to show above, List's argument using the natural ontological attitude delivers an ontology of agential alternatives that is not necessarily the same as that in play in ordinary discourse and folk psychology. If there were in fact an incompatibilist condition on responsibility, List's argument would be unaffected. List and other compatibilists may well be correct that there is no such condition but the argument offered here does nothing to dispel incompatibilist worries. The account List gives therefore talks past the traditional free will problem.

A referee for this essay recommended that I provide a simple example where the sense of alternative possibility List outlines would not be sufficient for our responsibility practices. In response, my criticism thus far is not dependent on my being able to show that. Instead, I have argued that whether or not incompatibilist conditions are in fact required, List's argument would be unaffected. Given that the free will problem hangs on whether such conditions are required, this amounts to a problem for List's account. However, below, I provide the physical location argument to show that List's sense of alternate possibility is not coextensive with the ordinary understanding of alternate possibility. I support the key premise in the argument with the atomic slit case. I offer this as a simple example of the kind asked for, as the ordinary notion is the one at play in moral discourse about moral responsibility.

3. Microphysical Information Is Relevant to Assessing Claims about What Agents Can Do

Above, I offer an immanent critique of List's position by showing that the kind of free will and alternate possibilities he generates might not be what we need given our responsibility practices. Below I argue for incompatibilism about alternative possibilities and physical determinism and offer a new line of criticism concerning the relations between the physical and agential modalities. This brings into focus the incompatibilist concern: it is surely natural to think that when I make claims about what I can do at the agential level, on the assumption that I am a physicalist, I am also committed to the supervenience bases (described in the microphysical frame of reference) of these different choices being physically possible. List denies this.

Someone might well object here that all this concern amounts to is that List's position is counterintuitive. And perhaps in pointing to this as a problem, I am simply demonstrating that I am still in the grip of the very modal confusion List was addressing. In particular, could this intuition be a product of the confusion when we talk in the same sentence across reference frames? Perhaps. However, I think it is possible to strengthen the objection that List's position is counterintuitive and argue that there is a substantial problem.

3.1 The 'Physical Location' Argument

Imagine that you are trying to decide whether or not to go and visit your friend in Manchester on Saturday. You live in London, and there is nothing that would

prevent this: you have the money to make the trip, you would be welcome, and you are healthy enough to travel. You can also stay in London if you so choose. At the coarse-grained agential frame of reference, there is nothing stopping you either way. It is Friday night, and you are wondering what to do in the morning, specifically wondering whether you should go or not. You think (correctly, according to List's model) that you *could* visit Manchester or you could stay in London. Focus on the modal claim you make on Friday night, the claim 'I could go to Manchester on Saturday'. Now suppose that physical determinism is true at your world, and you do not end up going, instead staying in London to work. According to List, the microphysical fact that your body's constitutive atoms were determined to be in that region of space we call London all weekend does *not* falsify the agential claim made on Friday night. But that claim now seems manifestly false. It seems that it does exactly that. How could you have gone to Manchester if the atoms that make up your body were always going to be in London on Saturday? Physical information about where your body's constitutive particles are going to be in spacetime is relevant to agential claims. You cannot (at least in this life) go somewhere without your body in tow. If this is correct, then it is not the case that determinism at the microphysical frame of reference is irrelevant to modal claims at the agential frame. My claim is this is what we are committed to in our ordinary practice. It is part of the ordinary understanding of ourselves and our place in the world that we cannot go somewhere without our constitutive particles in tow, but crucially where *they* are is constrained by the microphysical frame of reference—that is, the past physical states and the laws of physics.

The above argument is generalizable to any decision an agent is going to make where they think they are choosing from a range of options. This seems to constitute a strong case for saying that our ordinary claims about what it is possible for us to do are threatened by the truth of microphysical determinism. More formally:

- Premise 1: If determinism is true then the atoms that are constitutive of my body could not have been other than where they were in physical space at t_n .
- Premise 2: If the atoms that are constitutive of my body could not have been elsewhere in physical space at t_n then my body could not have been elsewhere at t_n .
- Premise 3: If my body could not have been elsewhere at t_n then I could not have done otherwise at t_n .
- Premise 4: Physical determinism is true.
- Conclusion: I could not do otherwise at t_n .

And so, we arrive at the incompatibilist conclusion that physical determinism would entail that I could never do otherwise than I in fact do. Am I still equivocating between the agential and physical frames here? It might be countered that it is one thing to say that the particles constitutive of my body cannot be in a certain part of space *where all of this is expressed in the language of microphysics*, but that does not mean that *I cannot go elsewhere* when we are just talking at the agential

level. The latter sentence is explicitly not to be assessed in terms of microphysical possibility according to List. Perhaps, after all, the very equivocation List was outlining is still implicit in my argument. In response, my argument is designed to show how the different frames of reference are mutually constraining on each other as a result of our ordinary commitments: specifically by what seems intuitively the correct way to assess the truth of propositions expressed at the higher agential level, when we are in possession of physical level information. As I go through the premises below, I show why they are not question begging against List's position and why they are supported.

Premise one is expressing a proposition wholly about the physical level. It is written in English prose, not in the technical language of microphysics, but it should be understood as referring to the very same atoms modeled by physical theory and as saying that the atoms cannot be anywhere else than they were *as expressed in the language of that physics*. As such, premise one is not controversial; it is just a statement expressing the simple physical consequence of determinism. Premise three is to be understood as expressing a sentence wholly at the agential level. That you cannot be some place without your body is an uncontroversial part of our basic commitment, barring any sci-fi, the afterlife, or Cartesian possibilities. As such, premises one and three each operate at their respective levels and state uncontroversial propositions.

Premise two is where the controversy is. Premise two ('If the atoms that are constitutive of my body could not have been elsewhere in physical space at t_n then my body could not have been elsewhere at t_n ') is the frame bridging material implication crucial to my argument. The antecedent is expressing a proposition that (just as with premise one) refers wholly to the microphysical level. The consequent makes a statement (just as with premise three) about the agential level. Why then is premise two not simply question begging against List given his set up? I concede that my premise two would be question begging if there was no independent support for thinking that premise was true. However, that is not the case as premise two is a part of our ordinary commitment about what science means for our relation to the world. To best illustrate this, it is instructive to consider cases where specific differences at the micro-level function as difference makers for the macro or agential level. Atomic slit experiments can provide clarity about the way information from the physical level 'bleeds up' and affects the truth of propositions describing the macro level.

3.2. The Atomic Slit Example: Making the Physical Level Information Salient

Imagine we have a small physically deterministic system where atoms fired at a slit can either pass through the slit or instead impact the surface around the slit. If the atom passes through, the machine, at the macro level, turns on a red light that scientists watching the process can observe and so they know the atom passed through. If, after the atom is released, it hits the surface, a green light comes on instead. Although the system is physically deterministic, for reasons of calibration and other micro-variables, sometimes the atom passes through and sometimes it

does not. In List's model, we have an indeterministic system at the macro level, with both red light events and green light events being equally possible. Now consider one particular firing in which an atom passes through and the red light comes on. We might wonder if it was possible that the green light could have come on instead. List will say yes, that was possible because there is an equivalence class at the micro level consistent with the same macro events up to the firing, where in that class some atoms then go on to pass through and some do not. (Figures 1 and 2 above illustrate this point.) For List, the fact that some do not underwrites the truth of the proposition 'the green light could have come on instead' at the higher level of description. But this seems intuitively the wrong way to describe things. We are talking about that specific atom and we know (with hindsight) that because of deterministic physics and the set-up, that particular atom could not have been elsewhere as described in physical space then. But that fact *just is* what decides whether the lights come on red or green. We know that the atom described in the physical frame could not have been at that part of space necessary to cause the green light to come on. And hence, it seems very hard with this information in hand, to continue to maintain that the green light could have come on instead.

Rather than by hindsight, we can deliver the same result by having the forward-looking physical level information instead. Let us stipulate that we also have an electron microscope or particle detector array that is capable of measuring the exact velocity of the atom post release. Once we have this information, we can immediately see whether the atom will pass through the slit or not. The details of how exactly the set up would allow us to infer that the atom will or will not pass through the slit could be various. Perhaps all the possible trajectories that allow the atom to pass through have been gathered by a process of trial and error over many observations. This is all possible without having to compute information at the micro level and then try to translate it into macro level information. We can simply see the correlations between various trajectories and which lights come on and build a data set accordingly. There is no reason why such a data set would not allow us to predict with perfect accuracy after a sufficient number of observations. What does having that information do to the truth of propositions at the macro level? Before the particle detector array comes along, we might have been happy to say for any particular atom firing, that the red or green light can light up, but that does not seem right now we have this new information. If the array reading tells us that after the release of the atom, it will pass through the slit, there is surely no further sense in maintaining the truth of the proposition, talking of this particular firing, 'the green light can light instead'. We might, of course, continue to understand such claims as just referring to the more general fact that sometimes the green light comes on and sometimes red. That would be innocent. But what seems very hard to continue to maintain is that this *very* atom could have caused the green light to come on. How could it, given what we now know? Again, if this is true then we have the kind of information bleed between levels that List denies. I claim that cases like this bring into focus that the modalities at the macro level are more sensitive to fine grain information from the physical level than List allows. Furthermore, it is hard to see how there could be a principled

difference between the simple physical set up here in the slit experiment and the complex microphysical machines we are when we make choices as agents.

Why exactly does the atomic slit case support premise two in the physical location argument? Because once we have information about the actual world supervenience base at the microphysical level, we can make an inference about the supervenient macro level phenomena. Once we know the atom will pass through the gap in the sheet of other atoms that we call the slit, we know the red light has to come on and not the green. Analogously, if (somehow) some observer outside the system came to have the information that the physical supervenience base of my staying in London on Saturday will be the case, they can infer I cannot go to Manchester instead. It does not matter whether we could ever in practice get that information because we cannot translate between levels. *If* we had it then it immediately constrains what we can say about the macro level. List claims that there are alternate possibilities at the macro level *because* there is an equivalence class of microphysical realizers of multiply realizable macro states. But this is equally true of the slit case, where the microphysical equivalence class contains atoms both passing through and not. If this fact underwrote the modal semantics at the macro level, we would not have this reaction to the slit case when we get the specific information about the atomic location, but we do. This is evidence that my premise two is independently supported rather than question begging against List. List can of course argue we should use his modal semantics and adopt a form of prescriptive compatibilism (I would readily agree with him on this point) but again, this is not the point at issue in the traditional diagnostic free will debate.

Perhaps it might be objected that my premise two is not really graspable because the antecedent is just shorthand for a technically expressed physical level statement and therefore it is hard, to have robust or reliable intuitions about the truth of the material implication. In other words, it is one thing, when talking at the macro level, to refer to atoms passing through slits and have intuitions about whether modal claims are true. But this is not the form the information would be in. Could it be that my premise two only seems intuitive because it is expressed in a shorthand form and that if faced with the raw physical level statements we would be flummoxed, perhaps not having any intuition about them and what they meant at all? It is no doubt true that the antecedent of premise two is information that would be technically expressed. However, I do not see how there is any salient difference *once we know what it means for the macro level*. How can we come to know what raw physically expressed information means? As I outlined above, we can come to learn which bits of space, expressed in the language of physics, are the bits we call the slit and the slit surround, and so on. So again, even without bridge laws between the levels of reality, we can build a translation table, by seeing what various physically expressed locations mean. Once we have built this translation table for the slit setup, when given purely physical level information about particle location (expressed in the purely physical way), we can consult the table and immediately make inferences at the macro level and whether the lights will come on green or red. This is evidence that premise two is supported even when we understand the antecedent as expressed in its technical form. Premise two is supported because we can demonstrate, via consideration of these slit cases,

that physically expressed location information would allow us to make an inference at the macro level. Technically expressed, physical level information can become meaningful to us via a translation table. If this is true of these simple systems, why should it be any different for more complex systems, like human beings? If List's equivalence class modal semantics for the macro level were correct, we would not have this reaction to these cases. These cases bring into sharp focus that, although we do not normally (ever) have access to this kind of physical level information, *if* we had it, it would constrain how we would be willing to describe the macro level. It is logically possible that an outside observer could come to have that information about the regions of space we call Manchester and London at the macro level and the physically expressed locations of our constituent atoms. A relevant translation table could fall into the lap of an observer of our world system. If that happened, they could infer that I could not in fact go to Manchester on Saturday, any more than the green light could come on instead of the red in the slit experiment.

Another possible response to my physical location argument here might be that I have made the same mistake that List (2019) accuses Van Inwagen of making in the consequence argument. That is to say that I have illicitly placed a proposition from one frame of reference within the scope of an operator that is not meant to quantify over propositions from that frame. Although I agree with List's criticism of Van Inwagen in this regard, my argument is not subject to this worry as it can be stated using only material implications, which, although they connect sentences from the agential frame on the one hand and the microphysical on the other, do not make the quantification mistake. In the consequence argument, p_0 describes the full state of the world micro-physically at some point in the past, p represents the particular action of an agent, l is a proposition describing the laws of fundamental physics and the modal operator N is such that Np means ' p is true and there is nothing anybody could have done to make it false.' List points out that certain premises in the consequence argument place agential level propositions within the scope of physical operators:

First, if we tried to formulate the consequence argument in fundamental physical terms, we would not express proposition p and the N operator adequately, because these belong to the agential level. Second, if we tried to formulate the argument in agential-level terms, we would not express propositions p_0 and l, \dots , adequately, because these belong to the fundamental physical level. And third, it is doubtful whether 'mixed' propositions such as Np_0 , Nl , $N(l \rightarrow p)$, and $N(p_0 \rightarrow (l \rightarrow p))$ are well-formed at all, because N and p are agential-level expressions, while p_0 and l are physical-level ones. In short, the consequence argument mixes two levels of description that do not go together. (List 2019: 5-6)

In summary, it is because those key propositions are not well formed that the consequence argument is not statable and hence it cannot be asserted. List makes clear that mixed propositions can be well formed and statable when they are

conjunctions, disjunctions or material implications relating propositions from one frame to propositions from another. It should be apparent that my ‘physical location argument’ is fully statable using only material implication. Premise one is a proposition that expresses the consequences of determinism at the physical level. Premise two certainly refers to both levels of description and it also asserts a relation between the levels, but it merely connects them using material implication. Hence, I do not quantify in the way List points out is not well formed. Premise three is an uncontroversial claim about the agential level. If the argument is unsound, it is not because of the issue concerning illicit quantification that List points to in the consequence argument.

4. Conclusion

Even granting List can demonstrate we possess free will in the sense of being able to choose from among agential possibilities, it is not clear, and it does not follow from List’s arguments making use of the naturalistic ontological attitude, that this sense of free will is *the sense at issue* in the traditional debate—that is, the sense that we standardly conceive of as the control condition relevant to moral responsibility. That is because the theories that necessitate the reality of a ‘power over alternatives’ do not try to model moral practice. If so, then List is talking past the traditional free problem with this approach. Further, our intuitive understanding of the relationship between ourselves described at the macro level and also as physical entities composed of particles that are the subject matter of physics is evidence that we cannot combine physical determinism with agential indeterminism in the way List contends. Physical determinism is still a threat.

JOHN DANIEL WRIGHT
INDEPENDENT SCHOLAR
john.2.wright@kcl.ac.uk

References

- Ayer, Alfred J. (1954) ‘Freedom and Necessity’. In *Philosophical Essays* (New York: St. Martin’s Press), 271–84.
- Berofsky, Bernard. (2005) ‘Ifs, Cans, and Fee Will: The Issues.’ In Robert Kane (ed.), *The Oxford Handbook of Free Will* (Oxford: Oxford University Press), 181–201.
- Berofsky, Bernard. (2012) *Nature’s Challenge to Free Will*. Oxford: Oxford University Press.
- Chisholm, Roderick M. (1964) ‘Human Freedom and the Self’. The Lindley Lecture for 1964, Lawrence: Department of Philosophy, University of Kansas.
- Clarke, Randolph. (2009) ‘Dispositions, Abilities to Act, and Free Will: The New Dispositionalism’. *Mind*, 118, 323–51.
- Elzein, Nadine, and Tuomas K. Pernu. (2017) ‘Supervenient Freedom and the Free Will Deadlock.’ *Disputatio*, 9, 219–43.
- Fara, Michael. (2008) ‘Masked Abilities and Compatibilism’. *Mind*, 117, 843–65.
- Fine, Arthur I. (1984) ‘The Natural Ontological Attitude’. In Jarret Leplin (ed.), *Scientific Realism* (Berkeley: University of California Press), 83–107.
- Frankfurt, Harry. (1969) ‘Alternate Possibilities and Moral Responsibility’. *Journal of Philosophy*, 66, 829–39.

- Hume, David. (1993) *An Enquiry Concerning Human Understanding*. Indianapolis: Hackett.
- Kratzer, Angelika. (1977) 'What 'Must' and 'Can' Must and Can Mean'. *Linguistics and Philosophy*, 1, 337–55.
- Lehrer, Keith. (1968) 'Cans without Ifs,' *Analysis* 29, 29–32.
- Lewis, David. (1981) 'Are We Free to Break the Laws?' *Theoria*, 47, 113–21.
- List, Christian. (2014) 'Free Will, Determinism, and the Possibility of Doing Otherwise'. *Noûs*, 48, 156–78.
- List, Christian. (2019) 'What's Wrong with the Consequence Argument: A Compatibilist Libertarian Response'. *Proceedings of the Aristotelian Society*, 119, 253–74.
- List, Christian. (2019) *Why Free Will Is Real*. Cambridge, MA: Harvard University Press.
- List, Christian, Gregg Caruso, and Cory Clark. (2020) 'Free Will: Real or Illusion. A Debate'. *The Philosopher*, 108, 61–80.
- Maier, John. (2015) 'The Agentive Modalities'. *Philosophy and Phenomenological Research* 90, 113–34.
- Moore, George E. (1903) *Principia Ethica*. Cambridge: Cambridge University Press.
- Smart, J. J. C. (1961) 'Free Will, Praise and Blame'. *Mind*, 70, 291–306.
- Smith, Michael. (1997) 'A Theory of Freedom and Responsibility'. In Garrett Cullity and Berys Gaut (eds.), *Ethics and Practical Reason* (Oxford: Oxford University Press), 293–317.
- Taylor, Christopher, and Daniel Dennett. (2005) 'Who's Afraid of determinism? Rethinking Causes and Possibilities.' In Robert Kane (ed.), *The Oxford Handbook of Free Will* (Oxford: Oxford University Press), 257–76.
- Quine, W. V. (1977) *Ontological Relativity and Other Essays*. New York: Columbia University Press.
- Van Inwagen, Peter. (1983) *Essay on Free Will*. Oxford: Oxford University Press
- Vihvelin, Kadri. (2004) 'Free Will Demystified: A Dispositional Account'. *Philosophical Topics*, 32, 427–50.
- Vihvelin, Kadri. (2013) *Causes, Laws, and Free Will: Why Determinism Doesn't Matter*. New York: Oxford University Press.
- Weir, Ralph. (2016) 'Relative Modality and the Ability to Do Otherwise,' *European Journal of Philosophy*, 12, 47–61.
- Wright, John Daniel. (2016) 'Varieties of Compatibilism: An Evaluation of Compatibilist Approaches to the Free Will Problem'. PhD diss. King's College London.