

The demand for contrastive explanations

Nadine Elzein¹ 

Published online: 22 February 2018

© The Author(s) 2018. This article is an open access publication

Abstract A “contrastive explanation” explains not only why some event *A* occurred, but why *A* occurred *as opposed to* some alternative event *B*. Some philosophers argue that agents could only be morally responsible for their choices if those choices have contrastive explanations, since they would otherwise be “luck infested”. Assuming that contrastive explanations cannot be offered for causally undetermined events, this requirement entails that no one could be held responsible for a causally undetermined choice. Such arguments challenge incompatibilism, since they entail that causal determinism is a prerequisite for moral responsibility. However, I argue that for a significant class of choices, even if we stipulate that they are determined, we will be unable to provide a *relevant* contrastive explanation. Hence causal determinism is no remedy for luck infestation, and compatibilists do not fare much better than incompatibilists in the face of this requirement. This should serve to weaken its philosophical appeal.

Keywords Compatibilism · Contrastive explanations · Free will · Incompatibilism · Indeterminism · The Luck Objection · Moral responsibility

1 Introduction

Some philosophers have argued that we cannot hold any agent morally responsible for a choice unless we are able to provide a contrastive explanation for that choice. A contrastive explanation, in this context, is one that causally explains why the agent chooses to perform a particular action *A as opposed to* some alternative action *B*. Contrastive explanations are taken to be necessary for moral responsibility,

✉ Nadine Elzein
nadine.elzein@univ.ac.uk

¹ University College Oxford, Oxford OX1 4BH, UK

typically, on the basis that they are thought to be crucial to control. If an agent's choice cannot be contrastively explained, then it is thought to be (in some crucial sense) inexplicable, and proponents of this objection maintain that an agent cannot be responsible for an inexplicable happening; such happenings would essentially be a matter of luck. Hence, the demand for contrastive explanations often forms the basis of a certain species of *Luck Objection* to incompatibilist accounts. Since it has often been supposed (quite independently of the free will dispute) that we could not offer contrastive explanations for causally undetermined events (Salmon 1971, 1984; Railton 1978, 1981; Lewis 1986a, b), it may seem as if indeterminism introduces luck of the sort that undermines, as opposed to furthering freedom.

This sort of argument has perhaps been most forcefully formulated by Mele (1999a, b, 2005, 2006), although it should be noted that Mele himself does not endorse the conclusion; he poses it as a puzzle for libertarian accounts, for which we might hope to develop solutions. In this way, he compares himself to a theist who explores the problem of evil, not with the aim of refuting theism, but with that of encouraging the development of theodicies (Mele 2017, pp. 170–172). Nonetheless, Mele's arguments have been influential. Almeida and Bernstein (2003) explicitly endorse the conclusion of the argument, and Haji (2000, 2001) endorses a closely related objection.

Since objections of this kind generally take such luck to be positively inconsistent with moral responsibility, it not only presents a problem for incompatibilists, who maintain that indeterminism is necessary for moral responsibility, but also for any compatibilist who wishes to maintain that neither determinism nor indeterminism threatens free will. However, it supports the claims of some compatibilists that determinism is a prerequisite for free will. The purpose of this discussion will be to outline a difficulty faced by anyone who (a) propounds versions of the Luck Objection that rest on the claim that contrastive explanations are required for moral responsibility, and (b) hopes that, so long as we are willing to concede the truth of causal determinism, we will be able to preserve most of our common-sense responsibility attributions. Proponents of such objections generally endorse both claims.

It should be noted that the species of the Luck Objection considered here, which appeals to a contrastive-explanations requirement, is just one among many. There are other species of Luck Objection, which I will not discuss. I will not say anything here about "Disappearing Agent" arguments (Van Inwagen 1983, 2000; Mele 1999a; Pereboom 2001, 2004, 2014; Haji 2000; Almeida and Bernstein 2003; see also Hobart 1934; Smart 1961). Likewise, I will not say anything about much older versions of the Luck Objection, which rest on the claim that undetermined actions would need to be uncaused, irrational, or both (Collins 1717; Hume 2000; original 1740, see also Hobart 1934). Nor will I discuss *other* problems associated with indeterminism and freedom, which don't rest on worries about luck, such as the problem of forking paths (Waller and Waller 2015), or the Mind Argument, as formulated by Van Inwagen (1983), where indeterminism is argued to be directly inconsistent with free choice. In this paper, I endeavour only to address versions of the objection that rest on a demand for contrastive explanations.

One way to counter such arguments is to directly attack the claim that contrastive explanations are relevant to control at all. I believe that there are successful arguments of this sort already in the literature (see especially Franklin 2011 and relatedly Mele 2006, 2017). There is also a significant controversy about whether contrastive explanations require determinism at all (Hitchcock 1996, 1999), and this can also be used to cast doubt on whether the demand for contrastive explanations can be used to ground any version of the Luck Objection (Franklin 2011).

I wish to outline a different sort of worry. Regardless of whether contrastive explanations require determinism, and regardless of whether such explanations can be relevant to moral responsibility, I will argue that the presence of determinism may not guarantee (or even render likely) that we can provide contrastive explanations that *are* relevant to moral responsibility, at least for a significant class of choices. This requirement may prove difficult to fulfil regardless of whether determinism obtains. Since such arguments are typically intended by their proponents to show that incompatibilist accounts are at a disadvantage in comparison to compatibilist ones, this would significantly undermine the philosophical appeal of such requirements.

2 The need for relevant contrastive explanations

Arguments that draw explicitly on a demand for contrastive explanations are given by Mele and Almeida and Bernstein. Haji does not appeal explicitly to contrastive explanations, but offers a closely related line of argument. Mele puts the point as follows:

In the actual world, Joe decides at t to A . In another world with the same laws of nature and the same past, he decides at t not to A . If there is nothing about Joe's powers, capacities, states of mind, moral character, and the like in either world that accounts for this difference, then the difference seems to be just a matter of luck (Mele 2005, p. 384).

While Mele does not, ultimately, suppose that this sort of "present luck" is necessarily destructive to moral responsibility and control (Mele 2006, p. 70, 2017, pp. 170–172), others have been less optimistic. e.g. Almeida and Bernstein, who direct their argument primarily at Kane's (1999) account, argue as follows:

One way to recognize that we should demand contrastive explanation in addition to plain explanation is that without this additional requirement we will be given an explanation of an act that is compatible with the unactualized alternative act. So, in principle, Kane may provide an explanation of why John resisted temptation that is consistent with John's submission to temptation. We know [...] why John did what he did but we are at a loss to know why John acted as he did to the exclusion of the other alternative open to him. To have control over an action does not merely demand that the agent have the power to perform that action; it should also demand that the agent have the power not

to perform an alternative act that he could have performed (Almeida and Bernstein 2003, p. 112).

On a similar note, Haji says the following:

Quite independently of whether one wants to side with incompatibilists or compatibilists over the free will issue, a reasonable hypothesis is that there must be *some* difference, other than a difference that is a function solely of indeterminism in the actional pathway of the disparate choices, that accounts for the difference in those choices, *if* there is to be a plausible reasons-explanation of either. Action theorists, then, may well want to insist that, given type-identical pasts, if Jones voluntarily and intentionally chooses in one way and his counterpart voluntarily and intentionally chooses in another, their choices are... a matter of luck. Further, they would have good reasons to insist that if the choices were *different and were not* the result of luck, then there would be an explanation – one that *need not* be *deterministic* – in terms of prior reasons or motives that would account for the difference in choices (Haji 2000, p. 222).

If Jones and Jones* have type-identical pasts, and Jones chooses to *A* while Jones* chooses to *B*, these theorists will maintain that we do not have the sort of explanation we need to ensure that Jones's choices are not luck-infested in such a way as to undermine Jones's responsibility.

While Haji supposes that the relevant sort of explanation needn't be deterministic, such a stipulation is of no help to most incompatibilists. If a given choice is causally undetermined, it will (arguably by definition) be possible for type-identical agents to make divergent choices from one another when their histories and the details of their present circumstances are identical. Since incompatibilists typically require indeterminism *at the time of choice*, as opposed to indeterminism that occurs at earlier times in an agent's history, such a possibility is clearly of little use to most incompatibilists.¹

One potential reason why Haji might suppose that indeterministic explanations may nonetheless, in theory, count as explanations of the right sort is that he is not merely defending the need for contrastive explanations, but defending the need for some prior *action* on the agent's part, which *ensures* that they perform the action they do in contrast to the alternative. As Franklin notes, however, the latter argument renders free choices incoherent for compatibilists and incompatibilists alike; if the agent must be able to do something in advance of a choice that ensures that it occurs, then we face a dilemma about the status of this prior action itself; if this prior action does not meet the same condition, it will not itself count as a free and responsible action, and hence it becomes mysterious why this should help to ground the freedom of the subsequent choice it ensures. In contrast, if we say that this prior action must *also* be free, then we require an infinity of free acts prior to

¹ While some have suggested that we may construct incompatibilist accounts that avoid the problem by locating indeterminism earlier in the agent's history (e.g. Dennett 1978; Mele 1995, 1999a, 2006; Fischer 1995; see also Ekstrom 2003) few incompatibilists defend this view.

anyone's acting in a way that counts as morally responsible. Clearly the barrier to meeting *that* condition is our finitude, as opposed to indeterminism (Franklin 2011).

A charitable reading of Haji's argument, then, places it in the same camp as Mele's and Almeida and Bernstein's, in appealing to a demand for contrastive explanations. Hence the problem will have to be as follows: If type identical agents can make divergent choices in identical circumstances, despite their identical histories, there can be no explanation for why either agent chooses as they actually do, as opposed to choosing the alternative, and *this* is what is taken to introduce luck, and hence is taken by some to be freedom undermining.

Such views are taken to be problematic for incompatibilists because it is thought that contrastive explanations of the right sort could not be provided for undetermined events. This assumption itself is contested.

Hitchcock notes that while the same events would be present in the history of an undetermined choice, regardless of which of the alternative outcomes result, those events would typically occupy different *causal roles* in relation to these different outcomes. This would potentially allow there to be contrastive explanations for undetermined events, so long as we allow a "different causes" reading, as opposed to a "different histories" reading (Hitchcock 1999, p. 593); on this model, an agent's reasons for choosing *A* would be causally active in one case, while an agent's reasons for choosing *B* would be causally active in the other, and we could cite the differing causal roles of these reasons as the basis of a contrastive explanation. However, such a move is blocked, if we follow Lipton (1991), and suppose that there is a Difference Condition on contrastive explanation, as follows:

DC: ...to explain why *P* rather than *Q*, we must cite a causal difference between *P* and *not-Q*, consisting of a cause of *P* and the absence of a corresponding event in the history of *not-Q*. (Lipton 1991, p. 43)

Lipton specifies that "a corresponding event is something that would bear the same relation to *Q* as the cause of *P* bears to *P*" (Lipton 1991, p. 44). Given that an agent has similarly strong reasons for choosing either *A* or *B*, the rival set of reasons will appear to constitute precisely such a "corresponding event".

Hitchcock notes, however, that which corresponding events would constitute an equivalent role in any rival "explanation" can shift depending on the prior assumptions and information available. That is because such factors make a difference with respect to which pieces of information are illuminating from a given agent's perspective, and this is, ultimately, the pragmatic point of explanations. This means that whether we have an adequate contrastive explanation may depend more heavily on the prior knowledge of the person to whom the explanation is offered than on whether the information offered pertains to deterministic or indeterministic causal factors (Hitchcock 1999).

Those who doubt that contrastive explanations can be offered for indeterministically caused choices must, then, accept certain presuppositions. Firstly, they must accept something along the lines of Lipton's Difference Condition, in order to rule out a "different causes" reading, and secondly, they must either not be using "explanation" in a sense that connects it to the pragmatic conditions under which some piece of information might count as illuminating or, alternatively, must

presuppose that the explanation is theoretically directed at someone who already has a complete knowledge of every detail of the agents concerned and their causal histories.

While there is a dispute to be had about the rationale and justifiability of these presuppositions, I won't explore these points any further. Let's instead grant the relevant presuppositions, and hence grant that determinism is a minimal requirement of offering a contrastive explanation of the right sort. On this account, a contrastive explanation is not *merely* one that "explains" why the agent does *A* as opposed to *B*, but one which cites causal factors sufficient to establish *both* that the agent does *A* and that the agent does *not* do *B*. This genuinely does seem to entail determinism (at least at the time that the choice is made), and hence the claim that contrastive explanations are required for moral responsibility, on this reading, does entail that indeterminism is positively inconsistent with moral responsibility.

Even if we grant this much, however, determinism alone will by no means guarantee us a *relevant* contrastive explanation. While there is a dispute to be had about whether determinism is necessary for the right sort of contrastive explanation to occur, it would be absurd to suppose that determinism is sufficient for it. That is because no one supposes that just *any* contrastive explanation will suffice to render an agent morally responsible.

Suppose that someone tampers with Niyah's brain, forcing her to choose *A*. This would certainly explain why she chose *A* as opposed to *B*. But this clearly would not give her any special control over her actions. We need our contrastive explanation to be of the right sort, if it is to be relevant to an agent's moral responsibility. Understandably then, Mele specifies that our explanation must have something to do with the agent's *powers, capacities, states of mind, moral character, and the like*, while Haji specifies that there must be some explanation *in terms of prior reasons or motives* for the agent's choice between alternatives. While Almeida and Bernstein are less clear about the conditions under which a contrastive explanation will be relevant to control, they generally describe the requirement in terms of *powers* of the agent. Presumably, any explanation that is supposed to establish a relevant "power" on the agent's part will have to appeal to relevant features of the agent's psychology; otherwise it would be hard to understand what is supposed to distinguish the exercise of a power from a mere happening.

All suppose then, either implicitly or explicitly, not only that contrastive explanations are needed in order to remove luck, but that such explanations must involve relevant features of the agent's psychology; their reasons, motives, character, state of mind, etc. Since the demand for relevant contrastive explanations will only be met insofar as our explanation draws on appropriate features of the agent's psychology, we need to examine the relation between determinism at a physical level and explanations at the level of psychology.

3 Physical determinism and psychological explanation

Suppose we live in a physically deterministic world. This does not imply that there are deterministic laws of psychology. In fact, phenomena at the psychological level typically elude deterministic explanation. It is widely conceded that we cannot identify “strict” laws governing psychology (or, for that matter, governing many of the special sciences).

Davidson famously argued that the mental realm is completely anomalous (Davidson 1970, 1974). On this model, it barely makes sense to even talk about laws of psychology. But even Davidson’s opponents, while insisting that the mental realm is not entirely anomalous, typically agree that there are no *deterministic* laws at the level of psychology (e.g. Lycan 1981). At the best, we might have hedged laws, which merely hold *ceteris paribus* (Fodor 1989).

The fact that there are no deterministic explanations available at the level of psychology prompts questions about the bearing of physical determinism on explanations of human agency. Some philosophers suppose not only that laws and explanations, but even *modalities* at the level of agency might be completely detached from what is possible at a physical level. On this basis, it has been argued that there may be genuine alternative possibilities at the level of agency, even if such alternatives are physically impossible given the history (see especially Kenny 1978; List 2014; List and Menzies 2017; though related views can be found in Taylor and Dennett 2002; Berofsky 2010, 2012; Roskies 2012; Ismael 2013, 2016).

I am highly doubtful about the idea that any agent may be granted the ability at a psychological level to do what is physically impossible in the circumstances, given determinism (Elzein and Pernu 2017). But the present argument does not require anywhere near as strong claim as this. It requires only the following much weaker claim: Even if we grant that the world is causally determined at the physical level, in the case of a significant class of decisions, we can rarely, if ever, offer deterministic causal explanations that rest on relevant features of an agent’s psychology.

The class of decisions that I have in mind are the difficult decisions; the ones that might be regarded as non-trivial, where *trivial* and *non-trivial* are defined as follows:

Non-Trivial Decision:

A decision in which the agent’s reasons for action do not speak decisively in favour of a single course of action. i.e. a difficult decision in which there is a conflict of motivations, desires, values, etc.

Trivial Decision:

A decision in which that agent’s reasons for action do speak decisively in favour of a single course of action. i.e. an easy decision in which there is no conflict of motivations, desires, values, etc.

Obviously, a decision could be trivial, by this definition, even if it concerns significant matters. Deciding to call an ambulance for a terribly injured loved one is an important matter, but in terms of the difficulty of the decision itself, there is no need to agonise and no difficulty involved in determining what to do. We may even

be doubtful about whether agents make “choices” in cases of this sort at all. The decision, in this sense, is trivial. Likewise, choosing between similarly appealing dessert options is hardly a life-and-death matter, but the decision itself could count as “non-trivial”, insofar as the agent’s reasons are similarly balanced enough for them to have difficulty choosing and a tendency to agonise.

Non-trivial decisions pose a problem for those who defend the demand for relevant contrastive explanations. The features of an agent’s psychology that seem relevant to moral responsibility (their characters, states of mind, reasons, motives, etc.) afford poor materials for accurately predicting agents’ non-trivial decisions. Such decisions are *not* the uniquely explicable outcome of relevant features of an agent’s psychology, regardless of whether there are deterministic explanations available for them at some level or other.

Even supposing there are, as Fodor suggests, laws of psychology that hold *ceteris paribus*, these will not produce useful predictions when an agent’s reasons conflict. Perhaps *all other things being equal*, if I want beer, I will order beer. But conflicts of motives occur precisely when other things are *not* equal. If I want beer, but I also want to stay sober, such *ceteris paribus* laws fail to yield good predictions (Elzein 2010). If my choice is guaranteed by deterministic physical laws, there will certainly be a contrastive explanation available at some level or other, but this does not establish that any contrastive explanation is available at the level that’s relevant to moral responsibility (even by the standards proposed by defenders of the demand themselves).

Perhaps if you knew enough about the microphysical structure of my brain, you could predict perfectly that one neural pathway would prevail over another, and hence that I will choose beer over sobriety. But this will not satisfy the requirement for a contrastive explanation that appeals to my psychological states. It would not be relevant to providing a plausible reasons-explanation.

While stipulating the truth of physical determinism will not provide a contrastive explanation of the right sort, what would provide such an explanation would be the stipulation that the psychology of *decision-making* operates via deterministic laws. This would require that in all cases, any agent’s reasons and motives must speak decisively in favour of a single course of action. This is not secured by physical determinism; rather, it requires us to live in a world that contains no *non-trivial* decisions.

In a discussion of Lipton’s Difference Condition, Almeida and Bernstein inadvertently highlight this very point. They give an example involving the receipt of similarly attractive job offers from different schools, A and B, in which they say the following:

Thus, we can explain why I go to A rather than B when the only job offer I receive is from A (this offer is the causal difference) while I cannot explain why I go to A rather than B when I receive equally attractive offers from both schools (the offer of B would bear the same relationship to my teaching at B that the offer from A plays in my teaching at B) (2003, p. 116).

The problem is that non-trivial decisions are commonplace, and would be commonplace *regardless* of whether determinism is true. If we take the requirement

for contrastive explanations seriously, all such decisions will be luck-infested, where this undermines moral responsibility. A world in which agents can typically meet the requirement would have to be a world in which no one ever receives equally attractive job offers from two different places; a world in which no one ever wants beer while they are trying to stay sober. Alternatively, it is a world in which morally responsible agents would have to freeze like Buridan's ass whenever they were faced with non-trivial choices.

As thus stated, the problem arises due to a disparity between the sorts of explanation available at a psychological level and the sorts of laws that exist at a physical level. Pérez de Calleja (2014) has also argued that compatibilists cannot easily escape this form of Luck Objection, but does *not* appeal to this disparity to make her case. She invokes the conditional analysis of the ability to do otherwise, which has traditionally been favoured by compatibilists. The relevant conditionals are understood in terms of what counterparts would have done in nearby possible worlds.

When agents are making non-trivial choices, there will usually be nearby possible worlds in which they decide differently; worlds with histories that differ in only insignificant ways from the actual world (i.e. in ways that have little rational bearing on the agent's reasons for action). Even if such worlds differ from the actual world at the level relevant to psychology, they may not differ in ways that speak to the agent's reasons. e.g. perhaps if the conversation had unfolded only a little differently (if the agent's interlocutor had paused, or phrased matters differently, say) she would have chosen otherwise. If this explanation is not one that rationalises the agent's choice, this may show that a *relevant* contrastive explanation is lacking, regardless of whether we suppose that determinism is true.

Psychologically identical agents can make divergent choices from one another, on Pérez de Calleja's view, even in deterministic settings. However, since this is a *conditional* reading of 'can', we can posit these abilities only by stipulating some antecedent difference in their histories or in the details of their present circumstances; identical agents could not make divergent choices from one another holding fixed the same history and circumstances. They could decide differently only in other nearby possible worlds. The rational landscape of these worlds may, however, be much like that of the actual world.

Defenders of the demand for contrastive explanations may object to allowing this conditional sort of ability to count. They typically stipulate that contrastive explanations will be unavailable in cases in which type-identical agents with *identical histories* are able to make divergent choices. It seems that this could only happen in an indeterministic setting.

This response may seem dubious, insofar as it requires us to suppose that all situational differences are relevant to an agent's rationality. But any plausibility this response has breaks down completely once we *also* see that the sorts of difference that furnish us with contrastive explanations often have no bearing on the agent's psychology at all. Haji and Mele explicitly state that relevant contrastive explanations must involve reasons, motives, psychological states, etc. While differences in circumstance, if they are located at a level that affects the agent's psychology, may be thought to inevitably be relevant to the agent's psychology,

differences located solely at the physical level do not obviously have a relevant bearing. In short, even if we hold laws, history, and present psychological states fixed, causal determinism is no remedy to luck infestation in the case of non-trivial choices.

There are, however, some potential responses to this argument. Firstly, we might drop the stipulation that relevant contrastive explanations would have to involve an agent's psychological states. Perhaps it's enough that there is the right sort of contrastive explanation available at a physical level. This would be secured by determinism. Alternatively, we might argue that while explanations at other levels are not relevant to an agent's control *in themselves*, their presence either directly implies or else constitutes evidence that there are *also* relevant contrastive explanations available at the psychological level. Let's examine these responses.

4 Will contrastive explanations at a neural level do?

Even if there are no contrastive explanations available that involve an agent's psychological states, perhaps it's enough that there is an explanation available at some level or other. Obviously, not just any contrastive explanation would do; we could not tolerate a theory according to which an agent may be rendered morally responsible purely on the basis that someone has manipulated her brain in some way, even if this *does* enable us to explain why she chose precisely as she did. We still need *some* restrictions on the sorts of contrastive explanations that can count as relevant. But perhaps a standard of relevance that restricts us to psychological factors, like reasons and motives, is unnecessarily stringent.

Perhaps a contrastive explanation that appeals to natural activity at a neural level will do. After all, events involving my brain chemistry or synapses may plausibly seem to be a part of my identity. We might therefore suppose that if we have an explanation for a choice that appeals to these features, this counts as an explanation that is relevant to my moral responsibility.

This strategy requires some substantial philosophical commitments. It requires us to suppose that agents can be identified with various sorts of brain activity, regardless of whether that activity has any special bearing on their psychology. Not everyone is likely to be sympathetic to such a highly reductive view.

Moreover, the supposition that any "natural brain activity" should be regarded as relevant one's moral responsibility, regardless of its bearing on one's psychology, is inherently implausible, at least without further qualifications. As stated, it does not exclude bodily activity caused by epileptic seizures, brain tumours, etc. from being counted as genuine agency. The obvious way to distinguish brain activity that is part of an agent's identity from brain activity that's external to it, is by appeal to the bearing of that activity on relevant features of an agent's psychology. If we rule this out from the start, it's hard to see how we might go about distinguishing the sorts of neural activity that constitutes agency from the sort which interferes with it.

The mere fact that our contrastive explanation appeals to an agent's brain activity seems insufficient to render it relevant to that agent's control. If a neural explanation is to be relevant, it must be *because* of its implications for the agent's psychology.

And if the demand for contrastive explanations is to be compelling, it must be because meeting such a demand rules out cases in which agents can choose differently from their *psychologically* identical (as well as merely *physically* identical) counterparts in exactly similar situations. This brings us to the second line of response.

5 Is determinism indicative of a relevant contrastive explanation?

Determinism may not guarantee that there is a relevant psychological difference between any pair of identical agents that can make divergent choices in the same circumstances, but perhaps we should suspect that there is always (or at least often) bound to be *some* relevant, psychologically situated, contrastive explanation present whenever there is one available at the physical level (although we might not know what it is). Contrastive explanations located at a physical level are not inherently doing any work, but they are at least *indicative* of a relevant contrastive explanation knocking around somewhere in the vicinity. Hence physical determinism typically renders it likely that agents' choices have relevant contrastive explanations.

There are two main problems with this suggestion. Firstly, there is no compelling reason to accept that the existence of a contrastive physical explanation implies, or even renders likely, that there is *also* a suitable contrastive psychological explanation knocking around somewhere. It's widely accepted that deterministic physical laws do not map onto the sorts of laws we might formulate in psychology or the other special sciences. Moreover, the thesis that any difference at the physical level implies a parallel difference at the level of psychology runs directly counter to the broadly accepted supposition that mental states are multiply realisable (Putnam 1967; Block and Fodor 1972; Fodor 1974; Block 1978; Pylyshyn 1984; Horgan 1993).

If the world is physically deterministic, there must be some deterministic mechanism via which the brain resolves the deliberative conflicts present in the case of non-trivial decision-making. But it's implausible to suppose that this must translate into any intelligible contrastive explanation we could formulate in psychological terms. Consider the way in which computers are programmed to decide between conflicting code paths. They typically utilise "random" number generators. But such devices usually operate entirely deterministically; they are not actually random. They are useful, nonetheless, because they produce outputs that *appear* arbitrary from any perspective that matters to us.

Presumably, if our brains are deterministic, they will also need to resolve conflicts with mechanisms of this sort. If that were the case, while there would often *be* contrastive explanations in neural terms, they would not correlate with any contrastive explanation that's relevant at a psychological level. Likewise, if minute differences at the level of synapses determine the direction in which we finally settle non-trivial choices, this would give us no obvious reason to suspect that these are indicative of any relevant psychological factors that determine the choice. In short, a deterministic physical explanation is no indication of a deterministic psychological one.

However, even if we accepted that the presence of a contrastive explanation at the physical level rendered it likely that there was some contrastive explanation at the psychological level present too (which we shouldn't), we would still be quite far from establishing that any such explanation is relevant. Not all aspects of psychology are relevant to an agent's control.

Recall, Haji requires that relevant contrastive explanations involve the agent's "prior reasons and motives" (2000, p. 222). Mele requires that they involve "powers, capacities, states of mind, moral character, and the like" (1999a, p. 280). Even if our physical explanation invariably had some bearing on the agent's psychology, it's not clear why we should assume it has a bearing on *these* features of it.

Suppose that Natalia must decide which of two equally talented candidates to offer a job to, and that her unconscious racial bias leads her to favour one of them, where this is the only factor that provides a contrastive explanation for her choice. Would this establish that she has more control over her choice than she would have had if she were free of the unconscious bias, and could have chosen either of them based on their strengths? Given that this factor does not provide any *reason* to favour this candidate, it's not obvious why it should afford her any extra control at all.

Admittedly, such unconscious biases could be regarded as part of Natalia's character. Given circumstances in which this is the only thing that prevents her from being able to decide either way, this could be the rationale for supposing that she could only be responsible insofar as she is driven by this bias; without it, nothing about her character would adequately explain her choice. However, the space of possible psychological factors that could theoretically explain Natalia's choice is broad. Some of these factors would be incredibly tangentially linked to her character.

Suppose Natalia had been subject to subliminal influence. One of the candidates wore a blue shirt, and she happened to have spotted something pleasing of that colour earlier in the day, unconsciously swaying her to feel slightly more positive towards this candidate. If this is *all* that swung the balance in favour of that candidate (and all that's stopping her from being able to choose either), it's not obvious that this helps to prevent the choice from being "luck infested". This explanation at least appeals to her psychology, but it's rather remotely connected to her reasons and motives, or her powers, capacities, states of mind, moral character, and the like. At best, it may relate to her "state of mind", but only at an unconscious level. Moreover, any number of arbitrary encounters are likely to have subliminal effects on an agent's psychology at any given time. And these encounters are themselves largely the result of blind luck. Why should this sort of luck be regarded as any less problematic?

6 Conclusion

There is no easy way in which to maintain the demand for contrastive explanations consistent with supposing that a compatibilist will easily be able to meet this demand where incompatibilists will fail. Causal determinism does not help us to provide a relevant contrastive explanation for non-trivial choices. If contrastive explanations of any plausibly relevant sort are required to prevent our choices from being luck infested, then determinism is no remedy to luck infestation.

An obvious solution to this problem would be to reject the requirement for contrastive explanations. Alternatively, it is open to compatibilists who accept the demand to embrace the conclusion that agents could never be responsible for non-trivial choices. However, since compatibilists are typically keen to endorse our common-sense responsibility attributions, this approach is likely to be unappealing. Insofar as compatibilists do wish to preserve those common-sense judgements, the arguments presented here should serve to significantly weaken the philosophical appeal of the demand for contrastive explanations.

Acknowledgements I am grateful to Professor Mark Kalderon, Michael Markunas, Dr Tuomas K. “Dick” Pernu, Professor Matt Soteriou, Dr Robyn Repko Waller, and other attendees of the London Mind Group, for valuable discussion and feedback in developing this paper, as well as to anonymous reviewers for their incredibly helpful comments.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Almeida, M., & Bernstein, M. (2003). Lucky libertarianism. *Philosophical Studies*, 113(2), 93–119.
- Berofsky, B. (2010). Free will and the mind-body problem. *Australasian Journal of Philosophy*, 88(1), 1–19.
- Berofsky, B. (2012). *Nature’s challenge to free will*. Oxford: Oxford University Press.
- Block, N. (1978). Troubles with functionalism. In C. W. Savage (Ed.), *Perception and cognition: Issues in the foundations of psychology. Minnesota studies in the philosophy of science* (Vol. 9, pp. 261–325). Minneapolis: University of Minnesota Press.
- Block, N., & Fodor, J. (1972). What psychological states are not. *Philosophical Review*, 81, 159–181.
- Collins, A. (1717). *A philosophical inquiry concerning human liberty*. London: Printed for R. Robinson, at Golden Lion in St Paul’s Church Yard.
- Davidson, D. (1970). *Mental events*, reprinted in his *Essays on actions and events*, pp. 202–228. New York: Oxford Clarendon Press, 1980.
- Davidson, D. (1974) Psychology as philosophy, reprinted in his *Essays on actions and events*, pp. 229–238. New York: Oxford Clarendon Press, 1980.
- Dennett, D. (1978). On giving libertarians what they say they want. *Brain storms: Philosophical essays on mind and psychology* (pp. 286–299). Montgomery: Bradford Books.
- Ekstrom, L. (2003). Free will, chance, and mystery. *Philosophical Studies*, 113, 153–180.
- Elzein, N. (2010). Conflicting reasons and freedom of the will. *Proceedings of the Aristotelian Society (Hardback)*, 110(3pt3), 399–407.

- Elzein, N., & Pernu, T. K. (2017). Supervenient freedom and the free will deadlock. *Disputatio*, *IX*(45), 219–243.
- Fischer, J. M. (1995). Libertarianism and avoidability: A reply to Widerker. *Faith and Philosophy*, *12*(1), 119–125.
- Fodor, J. (1974). Special sciences: Or the disunity of science as a working hypothesis. *Synthese*, *vo.*, *28*, 97–115.
- Fodor, J. A. (1989). Making mind matter more. *Philosophical Topics*, *17*(1), 59–79.
- Franklin, C. E. (2011). Farewell to the luck (and Mind) argument. *Philosophical Studies*, *156*(2), 199–230.
- Haji, I. (2000). Indeterminism, explanation, and luck. *The Journal of Ethics*, *4*(3), 211–235.
- Haji, I. (2001). Control conundrums: Modest libertarianism, responsibility, and explanation. *Pacific Philosophical Quarterly*, *82*, 178–200.
- Hitchcock, C. R. (1996). The role of contrast in causal and explanatory claims. *Synthese*, *107*(3), 395–419.
- Hitchcock, C. (1999). Contrastive explanation and the demons of determinism. *British Journal of the Philosophy of Science*, *50*, 585–612.
- Hobart, R. E. (1934). Free will as involving determinism and inconceivable without it. *Mind*, *43*, 1–27.
- Horgan, T. (1993). Nonreductive materialism and the explanatory autonomy of psychology. In S. J. Wagner & R. Warner (Eds.), *Naturalism: A critical appraisal* (pp. 295–320). Notre Dame, IN: University of Notre Dame Press.
- Hume, D. (2000). In Norton D. F. & M. J. Norton (Eds.), *A Treatise of Human Nature*. Oxford: Oxford University Press. (Original publication: 1740).
- Ismael, J. (2013). Causation, free will, and naturalism. In D. Ross & J. Ladyman (Eds.), *Scientific metaphysics* (pp. 208–236). Oxford: Oxford University Press.
- Ismael, J. (2016). *How physics makes us free*. New York: Oxford University Press.
- Kane, R. (1999). Responsibility, luck, and chance: Reflections on free will and indeterminism. *The Journal of Philosophy*, *96*(5), 217–240.
- Kenny, A. (1978). *Freewill and responsibility*. London: Routledge and Kegan Paul.
- Lewis, D. K. (1986a). Causal explanation. In D. K. Lewis (Ed.), *Philosophical papers* (Vol. II, pp. 214–240). Oxford: Oxford University Press.
- Lewis, D. K. (1986b). Chancy causation (Postscripts to “causation”). In D. K. Lewis (Ed.), *Philosophical papers* (Vol. II, pp. 175–184). Oxford: Oxford University Press.
- Lipton, P. (1991). *Inference to the best explanation*. London and New York: Routledge.
- List, C. (2014). Free will, determinism, and the possibility of doing otherwise. *Noûs*, *48*(1), 156–178.
- List, C., & Menzies, P. (2017). My brain made me do it: The exclusion argument against free will, and what’s wrong with. In H. Beebe, C. Hitchcock, & P. Price (Eds.), *Making a difference*. Oxford: Oxford University Press.
- Lycan, W. G. (1981). Psychological laws. *Philosophical Topics*, *12*(1), 9–38.
- Mele, A. R. (1995). *Autonomous agents: From self-control to autonomy*. Oxford: Oxford University Press.
- Mele, A. R. (1999a). Ultimate responsibility and dumb luck. *Social Philosophy and Policy*, *16*(2), 274–293.
- Mele, A. R. (1999b). Kane, luck, and the significance of free will. *Philosophical Explorations*, *2*(2), 96–104.
- Mele, A. R. (2005). Libertarianism, luck, and control. *Pacific Philosophical Quarterly*, *86*(3), 381–407.
- Mele, A. R. (2006). *Free will and luck*. New York: Oxford University Press.
- Mele, A. R. (2017). *Aspects of agency: Decisions, abilities, explanations, and free will*. New York: Oxford University Press.
- Pereboom, D. (2001). *Living without free will*. New York: Cambridge University Press.
- Pereboom, D. (2004). Is our conception of agent causation coherent? *Philosophical Topics*, *32*, 275–286.
- Pereboom, D. (2014). *Free will, agency, and meaning in life*. Oxford: Oxford University Press.
- Pérez de Calleja, M. (2014). Cross-world luck at the time of decision is a problem for compatibilists as well. *Philosophical Explorations*, *17*(2), 112–125.
- Putnam, H. (1967). Psychological predicates. In W. H. Capitan & D. D. Merrill (Eds.), *Art, mind, and religion* (pp. 37–48). Pittsburgh: University of Pittsburgh Press.
- Pylshyn, Z. (1984). *Computation and cognition*. Cambridge, MA: MIT Press.
- Railton, P. (1978). A deductive-nomological model of probabilistic explanation. *Philosophy of Science*, *45*, 206–226.

- Railton, P. (1981). Probability, explanation, and information. *Synthese*, 48(2), 233–256.
- Roskies, A. L. (2012). Don't panic: Self-authorship without obscure metaphysics. *Philosophical Perspectives*, 26, 233–342.
- Salmon, W. C. (1971). Statistical explanation. In W. C. Salmon et al. (Eds.), *Statistical explanation and statistical relevance* (pp. 29–87). Pittsburgh: University of Pittsburgh Press.
- Salmon, W. C. (1984). *Scientific explanation and the causal structure of the world*. Princeton: Princeton University Press.
- Smart, J. J. C. (1961). Free will, praise and blame. *Mind*, 70(279), 291–306.
- Taylor, C., & Dennett, D. (2002). Who's afraid of determinism? Rethinking causes and possibilities. In R. Kane (Ed.), *The Oxford handbook of free will* (pp. 257–277). New York: Oxford University Press.
- Van Inwagen, P. (1983). *An essay on free will*. Oxford: Oxford University Press.
- Van Inwagen, P. (2000). Free will remains a mystery: The eighth philosophical perspectives lecture. *Philosophical Perspectives*, 14, 1–19.
- Waller, R. R., & Waller, R. L. (2015). Forking paths and freedom: A challenge to libertarian accounts of free will. *Philosophia*, 43(4), 1199–1212.