



Friend, T. T. (2021). Why I'm not a Humean. *Pacific Philosophical Quarterly*. <https://doi.org/10.1111/papq.12398>

Publisher's PDF, also known as Version of record

License (if available):
CC BY

Link to published version (if available):
[10.1111/papq.12398](https://doi.org/10.1111/papq.12398)

[Link to publication record in Explore Bristol Research](#)
PDF-document

This is the final published version of the article (version of record). It first appeared online via Wiley at <https://doi.org/10.1111/papq.12398> Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available: <http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>

WHY I'M NOT A HUMEAN

BY

TOBY FRIEND

Abstract: There is an inconsistency between the access we have to our conscious lives and the Humean thesis of causal generalism. This was first drawn attention to by John Hawthorne, whose argument withstands a number of objections. Nevertheless, it has weaknessess. The first premise must be weakened if Humeans are to be compelled to accept it, and consequently, the second premise will have to be stronger to retain validity. I shore up the case against Humeanism by providing revised premises along with new defences of them. I show why this also provides a lesson for non-Humeans about the epistemology of causal relations.

1. *Introduction*

Humeanism is false and straightforwardly so, or so John Hawthorne has argued. His ‘argument from consciousness’, as I’ll refer to it, draws on the observation that Humeanism is committed to consciousness being dependent on events going on beyond what we should think it depends on. Here, it is verbatim.

- P1. An intrinsic duplicate of any region wholly containing me will contain a being with my conscious life.
- P2. There are causal requirements on my conscious life.

Therefore, Humeanism is false.

Under Hawthorne’s authorship, the possessives obviously refer to him. Under mine, they refer to me. In either case, the argument should have just the same force, so long as one is willing to allow that this author does indeed enjoy a conscious life much as Hawthorne does. (Alternatively, the reader may wish to interpret the possessives as referring to themselves to

Pacific Philosophical Quarterly •• (2021) ••••• DOI: 10.1111/papq.12398

© 2021 The Authors

Pacific Philosophical Quarterly published by University of Southern California and John Wiley & Sons Ltd. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

circumvent sceptical issues; certainly some arguments below will benefit from such reinterpretation.) I won't bother to ascertain whether the argument from consciousness is ambiguous between two homophonous arguments or two interpretations of the same argument.¹

The conclusion is supposed to follow because Humeans endorse a thesis of extrinsicity about causal relations.

EXTRINSICALITY. The causal facts pertaining to any subregion of the world are extrinsic to that region, supervening on the global distribution of freely recombinable fundamental properties.

Thus, the Humean should not accept that any intrinsic duplicate of any feature of a subregion (or, which I assume is also implied by the thesis, of any feature of an occupant of a subregion) of the world, which depends on causal facts pertaining to that region, will preserve those characteristics. Some recombinations of the global distribution of fundamental properties will preserve the intrinsic properties of any subregion but, because causal facts pertaining to that region are extrinsic to it, will not maintain features dependent on such causal facts. As I will sometimes put things, this is because events outside the region may *undermine* the relevant causal facts holding. Granting that my consciousness is dependent on causal relations, as per P2, Humeans must therefore deny that all intrinsic duplicates of regions wholly containing me will contain a being with the same conscious life, contrary to P1. So Humeanism (or one of the premises) is false.

For such a substantial conclusion, and initially plausible premises, Hawthorne's argument should certainly have been more widely discussed. Note in particular that the argument assumes nothing about the realisers of consciousness – they might be purely physical or involve some fundamental mental substrate. Nor does it assume much that is controversial about having a conscious life – it doesn't, for example, require the existence of phenomenal consciousness or qualia. Indeed, there is little up front in argument for the Humean to complain about, and whether they know it or not, they need a response. So why has there been so little traction? One reason is that Hawthorne's argument is not widely known.² Another is that the argument is known but believed to be easily defeated. Either way, my aim here is to show that while Humeans may capitalise on certain weaknesses in Hawthorne's specific formulation, a related argument can be offered, which is harder to undermine. Humeans really do seem to be, as Hawthorne put it, 'out of their minds'.

Here's the plan. Section 2 covers some objections that make no special mention of consciousness including one from Brian Weatherson, who has provided the only focused critique of Hawthorne's argument in print that I know of. Each of these objections have fairly straightforward responses. Hawthorne's argument does, however, have real weak points. In Section 3,

I explain why the bedrock intuition with which Hawthorne maintains P1 is insufficient in the face of countervailing intuitions and extant views that openly deny it. Moreover, I suggest that Humeans may be likely to accept stranger-sounding consequences about consciousness before they consider their position finally refuted. In response, I advise weakening the premise significantly to accommodate a looser connection between my (alternatively, Hawthorne's, your, or whoever's) conscious life and those of my intrinsic duplicates. A consequence of this is that the second premise must take up the extra slack. It's not obvious that Hawthorne's original defences of P2 can justify any stronger thesis, nor even that they are particularly successful at justifying the original premise. In Section 4, I therefore supply a new argument for a more committed second premise. Section 5 considers whether there may be an issue of consistency between the newly proffered arguments for both premises. I claim that there is not, but in doing so show that the non-Humean must reject the concession to the Humean that causal and nomic relations never make for an observational difference. Drawing these revised premises together in Section 6, I show how they serve to entail the same conclusion as Hawthorne's original and, therefore, why I'm not a Humean.

2. Preliminary objections and responses

In order to avoid the conclusion from the argument from consciousness Humeans might consider rejecting aspects of EXTRINSICALITY. The thesis is a consequence of two claims popular among Humeans. One claim accounts for the phrase concerning 'freely recombinable fundamental properties'. It states that every matter of fact (e.g. an event, state of affairs or property-instance) can be freely recombined with any other to make a possible world.³ Still more precisely:

RECOMBINATION. Anything can coexist with anything else so long as they occupy distinct spatiotemporal positions.

The other claim accounts for the extrinsicness of causal relations due to Humeans' commitment to generalism about causal and nomic relations:

GENERALISM. Necessarily, causal-nomic (c-n) relations are instances of universal regularities

For our purposes, a c-n relation can be understood as a relation instanced by a pair (e.g. a cause and effect) or larger n-tuple (e.g. the masses, distances and gravitational forces experienced by a system of celestial bodies) of events (although again, states of affairs or property-instances would also suffice) that count as an instance of causation or a law of nature. I will understand

universal regularities to be conjunctions of properties that hold throughout the entire universe (the schematic generalisations, *all Fs are Gs*, and *for all Fs, there is a G*, are indicative). The properties involved in universal regularities (e.g. those replacing F and G) are to be *purely qualitative*, meaning that they are not essentially picked out by singular terms. Universal regularities need not be so general as to fit the title ‘law of nature’, although they may well be *covered* by some combination of them. What is crucial is that the status of some instance of a c-n relation (i.e. some F and G events being related causally or nomically) counting as a c-n relation is dependent on the truth of some regularity involving F and G. Hence, the instance will not be intrinsic to its relata, because the global regularities could be different while intrinsic features remain the same.

Could Humeans object to the argument from consciousness by abandoning either RECOMBINATION or GENERALISM? Both options would seem to go against Hume’s own commitments. But Humeanism has come to indicate more an allegiance to David Lewis than Hume himself, and while Lewis defended both claims his doctrine of ‘Humean supervenience’, to which most references to ‘Humeanism’ today refer, is really only a commitment to the former claim. Someone might therefore fairly claim to be a ‘Humean singularist’, endorsing RECOMBINATION but denying GENERALISM (cf. Ducasse, 1926; Wilson, 2009).

In principle, I think Humean singularism deserves the name ‘Humean’. It also fails to be falsified by either Hawthorne’s argument from consciousness or the revised argument I formulate below. Nevertheless, given how rarely defended the position is it is understandable that Hawthorne didn’t take it into account. From hereon, I’ll assume the issue at hand is whether Humean generalism (i.e. an endorsement of both RECOMBINATION and GENERALISM) can be falsified by considering the causal dependencies of conscious humans, and that ‘Humean’ should be taken as an abbreviation of ‘Humean generalist’. To clarify, then, my aim in this article is only really to say way I’m not a Humean *generalist* (my apologies if the brevity of the title unfairly piqued your interest).

One will notice that GENERALISM is a thesis about both causal and nomic relations, whereas Hawthorne’s argument concerns only causal relations. Today, there are plenty of Humean non-believers in causal relations (cf. Price and Corry, 2007). It would be wrong, however, to think that this is a route out of Hawthorne’s refutation. After all, the non-believers typically hold that there are genuine *nomic* relations, and being Humean, nomic relations will just be instances of universal regularities. Hence, the argument is easily phrased in causal and/or nomic terms (this is how I’ll proceed in later sections).

A different objection has been made explicit by Brian Weatherston (2007) who suggests that the argument from consciousness is in fact invalid. Humeans, he claims, can avoid the conclusion because GENERALISM is

in fact consistent with allowing that c-n facts depend on specific regions after all. Humeans invariably think the laws are theorems of a maximally pithy deductive systematisations of the world (i.e. ones that trade-off simplicity and comprehensiveness or 'strength'; Lewis, 1973, 1983, 1994; Loewer, 1996, 2007; Braddon-Mitchell, 2001; Cohen and Callender, 2009). It is entirely coherent, then, that the laws could be region-relative, if the most systematic generalisations over all do better if they reference the ongoing in specific systems. As Weatherson observes, 'if the local regularities [of some specific region] deviate too far from the global regularities, then Humeans can and should say that different nomic claims (and hence causal claims) are true in this part of the world to the rest of the universe' (p. 531). Granting this, Humeans will expect there to be worlds where certain regularities L are maintained within a region r occupied by the intrinsic duplicate of a conscious human but breakdown beyond r in such a way that a theorem of the maximally pithy system for the intrinsic duplicate's world may be of the form 'In r , L ' (p. 532). And Weatherson claims that it remains logically possible that the laws concerning intrinsic duplicates of conscious persons are always like this and if so, the argument from consciousness is invalid (p. 533).

Even if he has shown the argument from consciousness to be technically invalid, Weatherson's observations aren't enough to show that the conclusion doesn't *as a matter of metaphysical necessity* follow from the premises. And indeed it is quite reasonable to suspect that there are worlds where the regularities beyond intrinsic duplicates of r will require that L is not specified in a law. This has, to my mind, been convincingly argued for by Pallies (2019). One way Pallies shows this is by considering 'Plenitude world', a world in which the following is true.

PLENITUDE. For every abstract pattern of fundamental property and relation instantiations, there is some set of objects which realizes that pattern.

He then observes that,

PLENITUDE is an extremely simple and informative description of the goings-on of Plenitude World. It's far simpler than a description which lists the patterns pertaining to various regions within Plenitude World. So, on the best-systems theory, we get a familiar result: the global pattern, described by plenitude, defeats the local patterns. PLENITUDE turns out to be a law. (p. 144)

Supposing that the causal relations in Plenitude world follow from whatever regularities can be gleaned from PLENITUDE, it is fair to suppose that an intrinsic duplicate in Plenitude world of a conscious human will fail to have the causal relations on which consciousness supposedly depends. The possibility of such a world is all that is required for the argument from consciousness to be sufficiently tight to warrant the conclusion.

Plenitude world is, of course, an extreme sort of a world, and Humeans might reasonably question whether it is a genuine possibility.⁴ The Humean must question more than this, however, if they are to make it plausible that in no world with an intrinsic duplicate of a conscious human is the best deductive system one without region-specific laws. Even if Plenitude world is too extreme, it seems reasonable to think that the Humean's adherence to RECOMBINATION will allow for some kinds of worlds whose laws are global enough to prevent an intrinsic duplicate of some actually conscious human having the causal relations on which their consciousness supposedly depends.

Anyway, Pallies also provides a different method of undermining the idea that duplicates of subregions must exhibit the same c-n relations (refer to note 10, p. 147). He points out that the c-n relations of one region *R* may not be present in an intrinsic duplicate region *R'* if the laws for *R'* are such as to prioritise the causal-nomic influence of some external source present at *R'* but not *R* (the regions may be in different worlds governed by different laws or in the same world at a suitable distance). Pallies suggests that such laws would permit so-called 'trumping' pre-emption (Schaffer, 2000), with the additional external influence in *R'* trumping the causal effects of the duplicates of causes in *R*.⁵ In that case, duplication of my conscious life would not be sufficient for preservation of the c-n relations on which it supposedly depends because there is the possibility of pre-empting alternative causes.

Weatherson's criticism on the basis of a failure of validity is therefore not a convincing objection to the argument from consciousness. As we'll shortly see, a different criticism of his focuses on the plausibility of P1. It is here I think the Humean may hope for more progress.

3. *Weakening P1*

If *y* is a duplicate of *x*, it is intrinsically exactly similar to *x*. Consequently, *y* will have all the intrinsic qualitative properties *x* has. But *y* needn't have all the extrinsic properties *x* has. For instance, *x* may be an uncle, or next to a vending machine, while *y* may not be. Similarly, extrinsic relations may fail to be retained between duplications of their relata. If *a* is taller than *b*, then a duplicate of *a* will be taller than a duplicate of *b*. But if *a* is closer to a vending machine than *b*, the same may not be true of their duplicates.⁶

One question we may ask concerning duplication is whether c-n relations are intrinsic or extrinsic to their relata (Menzies, 1999). One reason to think they are intrinsic is that we can often tell a causal relationship via direct observation. If one has witnessed a thrown stone smash a window pane via direct contact, then one would have witnessed that a causal relationship took place. Witnessing any intrinsic duplicate of such a scenario would invoke

exactly the same certainty, so causal relations must be intrinsic, or so it can seem.

But the Humean will demur. Perhaps one can learn via direct observation that a causal relationship has taken place, much as one can tell from a few confirming instances that some regularity holds. But there are worlds (*pace* Weatherston) in which precisely the same intrinsic scenario takes place between the stone and the window and yet there is no causation. That's because c-n relations in general are necessarily instances of universal regularities, as per GENERALISM. If the required regularities don't hold in those other worlds, because of some things elsewhere that undermine them, then the c-n relations won't hold no matter how much it looks like they do.

In fact, the Humean is committed to an even stronger thesis about c-n relations. Not only will intrinsic duplicates of the relata of c-n relations fail to necessarily preserve their status as c-n relations, but for the Humean, *there will be no region wholly containing a c-n relation whose duplicates necessarily maintain that c-n relationship*. That's because even a duplicate of an entire world, our world for example, can appear as a part of a larger world whose more general regularities swamp those appearing in the duplicate region. These more general regularities could therefore render those which underlie the c-n relations in our world no longer existent. In this sense, for the Humean c-n relations are *maximally extrinsic*.

The strangeness of these epistemic consequences of Humeanism are often noted, but that's something Humeans have learned to stand by and rightfully complain other views lack clear alternatives for. Humeans also have a plausible justification for their position. Even singularists must admit that causal relations aren't strictly observable in the sense of an observable 'tie' or 'spark'. Anscombe (1971) admitted as much when she remarked that,

it is true that the apparent perception of such things may be only apparent: we may be deceived by false appearances. Hume presumably wants us to 'produce an instance' in which *efficacy* is related to sensation as *red* is. It is true that we can't do that; it is not *so* related to sensation (pp. 136–137).

It can therefore seem granted on all sides that there is nothing distinctive, observationally speaking, about the existence of a c-n relation within any region. We might further suppose that this explains, for example, why observational scientific studies to ascertain the likelihood of some c-n generalisation are always defeasible. These studies attempt to determine whether or not some c-n relation exists from data. But there could always be some data that have not been observed and that would undermine any positive c-n hypothesis.

If non-Humeans are to accept such a lack of observable difference, they will need an explanation of it. Assuming theirs is a view according to which c-n relations are intrinsic, non-Humeans may have to settle for the ad hoc

claim that c-n connections are simply not the sort of thing which can make for the kind of observable difference that other properties and relations can (as per Anscombe's quote).⁷ Coherent as it is, this can sound a rather weak explanation without further justification. After all, we can, for example, point to observable differences between cases differing only with respect to whether or not the taller-than relation is and isn't present – there may be pairs/perspectives where it's hard to tell, but the intrinsicness of the relation means the fact of the matter is nevertheless always 'before our eyes'. So why shouldn't things be the same with c-n relations?

On the face of it, Humeans have a much better explanation for a lack of observable difference between regions differing only in whether or not there are c-n relations: namely, those relations' imputed maximal extrinsicness. For the Humean, there can be no region of the world observation of which would permit one to have observed grounds sufficient for any c-n relation, because those grounds are always maximally extrinsic. That's not to say the Humean can't accept that we may be very well attuned to where and when c-n relations do appear in our environment, perhaps with only the slightest evidence. The fact that there is a lack of observable difference between regions differing only in whether or not there is a c-n relation present need not, for the Humean, mean that when there is a c-n relation present that is not something which can't be learned of through observation. But being reliably well attuned to a phenomenon, and so perhaps also coming to know of it by observation, is not the same as directly observing grounds sufficient for it. And for the Humean, even if one were to observe the whole world, one would not thereby have knowledge of sufficient grounds, because one would also need the knowledge (besides a serviceable theory of c-n relations) *that one had* observed the whole world. It seems unlikely that we are ever to have contact with the world in this way, and this would explain why there can be no observable difference between regions containing c-n relations and duplicate regions without such c-n relations. Moreover, the availability of this potentially superior explanation provides an abductive justification (or inference to the best explanation) for c-n relations' maximal extrinsicity. Or so the Humean might suggest.

Put this apparently favourable Humean assessment of c-n relations' epistemology to one-side. Another question concerning intrinsicness is whether someone's conscious life is something shared by their duplicates or not, or more specifically, whether a duplicate of any region wholly containing me will contain a being with the same conscious life as me.⁸ This is, in effect, to ask about the status of P1. Hawthorne thinks it is 'wholly obvious' that they must, or that it is 'at the very least [...] highly intuitive'. '[I]f I am in pain, then any region that is not occupied by someone who is in pain will fail to duplicate any region wholly containing me' (p. 352). If that's right, then if it can also be established that there are causal requirements on instances of such consciousness (i.e. P2) then the Humean is clearly in trouble.

But for all Hawthorne's strength of intuition, there are simply too many views on the philosophical market which conflict with it. Hawthorne himself mentions Dretske (1995), who claims that phenomenal character is dependent on evolutionary history, and Lewis (1986a), who claims that phenomenal experience – 'how it feels' – depends on the causal role of the realising state for that experience in an 'appropriate population'.⁹ Hawthorne dismisses these views on the basis of his bedrock intuition that P1 is correct, but although the views are perhaps niche, his assessment is surely too quick.

Weatherson voices a counter-intuition to Hawthorne's with a thought experiment. He imagines two conscious beings Tweedledee and Tweedledum. Tweedledee faces a perfectly visually, aurally, olfactorily, gastronomically, and so forth symmetrical scene and consequently has an overwhelming feeling of symmetry. Tweedledum is just like Tweedledee except for having a few extra sensors and brain cells that correspond to an empty part of Tweedledee's brain. These additions allow Tweedledum to be able to sense an asymmetrical magnetic field as well as the otherwise symmetrical scene. Consequently, Tweedledum does not have the overwhelming feeling of symmetry that Tweedledee has. But if such a story is at all coherent, it suggests that my conscious life is not only dependent on my intrinsic properties, like the brain matter I have, but also on my extrinsic properties, like the brain matter I *don't* have.

Aside from (arguably) extreme views and combating intuitions, today it is an entirely legitimate position in the philosophy of perception to think that the external objects of perception are in fact directly and essentially involved in those perceptual experiences. These 'naive' or 'disjunctivist' views about perception may not be particularly natural for the Humean to adopt, but in doing so, they may be used to block the acceptability of P1. The disjunctivist, for instance, thinks that 'real perceptions are characterised as instances of relations borne to worldly objects, instances of relations which could not exist if their relata did not' (Steward, 2011, p. 139). The disjunctivist therefore denies that any intrinsic duplicate of me having a veridical experience (say) will be having the same experience.

We should, I think, remain open to intuition guiding our views on matters of consciousness. But the variety of ways in which one might justifiably oppose P1, and so also Hawthorne's bedrock intuition, strongly suggest in this case either we find some alternative defence of P1 or we revise it. My preference is for the latter option. To this end, it's worth noting that none of the above-considered reasons for opposing P1 deny that an intrinsic duplicate of me will have *some* conscious life, even if it is not quite the same conscious life nor even sharing any common experiential element. A simple fix to P1 would therefore be to change it to the following.

P1'. An intrinsic duplicate of any region wholly containing me will contain a being with a conscious life.¹⁰

P1' permits my intrinsic duplicates (in duplicate regions) to differ in their exact conscious state but asserts that they must have some conscious life given that I do. Such intrinsic duplicates might be those with different evolutionary histories to me or within different relative populations, or they might be related to me like Weatherson's Tweedledum is to Tweedledee, or as two intrinsically identical people are such that one is having a veridical experience and the other is having a hallucination. In all these cases, the duplicates differ from me in what conscious state they're in, but they don't differ from me in whether or not they have some conscious life.

The Humean might still protest, however, that there is no justification for the assumption that my duplicates will contain some conscious life. Perhaps as a consequence of external factors, some duplicates in duplicate regions need not have conscious lives at all. We could, I think, reasonably object to this suggestion on the grounds that it departs significantly from any of the considered views which take one's particular conscious state to be a partly extrinsic matter. Nevertheless, by weakening the first premise even further, we can find something the denial of which would really be undesirable for the Humean to accept. For even if we can tolerate the idea that so much as having a conscious life is dependent on circumstances going on beyond my peripheries, there must surely be a limit to what can be included; or, to put this another way, my having a conscious life cannot be maximally extrinsic.

We noted earlier that for the Humean, c-n relations are maximally extrinsic in the sense that there is no region duplication of which ensures that c-n relations are maintained within it. We granted that the Humean may have some abductive evidence for this hypothesis: that there do not seem to be regions of any size observation of which reveals anything distinctively causal or nomic. A region in which all the local matters of fact occurred by chance without c-n connections or with different c-n connections could look just the same to us. Or so it can seem. The Humean explains this by pointing out that knowledge of sufficient grounds for c-n relations requires knowing all the local matters of fact of the world plus the fact that those are all the facts. But the Humean can surely not give the same story about my having a conscious life. Unlike c-n relations, for which it is generally admitted exhibit of no observable tell-tale 'tie', it is just very hard to imagine that my being conscious is not observationally different, from my perspective, from whatever circumstances would have to be like for me not to be conscious. In general, there *is* a plain observable difference between being conscious and not being conscious in that in the former case one can be observing anything at all but in the latter case one is not observing anything.

It's crucial to note the triviality of this last point. In particular, it in no way commits one to any particular view about the nature of (my) consciousness. An illusionist will say, for example, that there is no technical difference between the conscious life of a human and a philosophical zombie that

duplicates only whatever physical (non-fundamentally psychic) stuff is relevant to the human's conscious life – both lack *phenomenal consciousness*. Nevertheless, illusionists don't deny that we have an 'inner life'.

Having the kind of inner life we have, they will say, consists in having a form of introspective self-awareness that creates the illusion of a rich phenomenology [...] Illusionists can say that one's experiences are like something if one is aware of them in a functional sense, courtesy of introspective representational mechanisms. (Frankish, 2016)

The ongoing debate over the phenomenal status of consciousness, and whether physical realisers are sufficient for consciousness, is entirely orthogonal to what I'm suggesting everyone, including the Humean, must grant. For even the most deflationary views about consciousness will allow that differences in one's conscious – 'inner' – life are directly observable; specifically they will be directly observable by the people whose consciousness is under consideration. For the Humean, at least, that must make the instantiation of consciousness rather different, epistemically speaking, from causal relations.¹¹

This all means that there isn't the same abductive support for the view that my having a conscious life is maximally extrinsic as there is for the view that c-n relations are. The datum to be explained that c-n relations don't seem to admit of an observable difference simply isn't paralleled by the case of consciousness, and so there is no justification for positing consciousness's maximal extrinsicness. Of course, we need to be careful in saying what exactly the issue is here. It remains the case for the Humean, as much as anyone, that c-n relations are 'observable' in the sense that we can be reliably attuned to their presence by observing their relata. And perhaps Humeans will hope the same for consciousness as well. But the crucial point being made here is that one's own consciousness plainly has a more direct observational mark than c-n relations will do under a Humean view of them. While it may be consistent with what we observe that what looks like a region containing a c-n relation might in fact not be, it just doesn't seem to make sense to say the same of consciousness. Indeed, it is not consistent with my observation of anything that the region which wholly contains me might not actually contain a conscious person.

In fact, the case against maximal extrinsicness is stronger than that. For if my having a conscious life were maximally extrinsic, this would mean that my having a conscious life is dependent not only on local matters of fact way beyond anything I am in causal contact with but also with the fact that the world is the size it is and no bigger. It is certainly bewildering to think that our conscious lives are so dependent. It would suggest that my ability to ascertain that I have an inner conscious life is achieved by putting me in contact with the totality of the world. Moreover, granting that my having a conscious life can be so dependent puts Humeans at odds with their very

own explanatory aims. As we've seen, the Humean will want to maintain that we should believe in c-n relations' maximal extrinsicness in part because that explains the lack of observable difference between the presence and absence of c-n relations. By parallel reasoning, if my having a conscious life is also maximally extrinsic, then it should also explain a lack of observable difference between the presence and absence of my conscious life. But it clearly doesn't explain this, because there *is* an observable difference. Therefore, if the Humean wants to maintain that my conscious life is maximally extrinsic they must abandon the claim that c-n relations' maximal extrinsicness is justified as an inference to a quality explanation of our poor epistemic contact with them.

In sum, there is good philosophical and even intuitive reason to question the bedrock intuition that supports P1. But these reasons nevertheless suggest the Humean should commit to a weaker premise (P1') that having *a* consciousness life is intrinsic. We considered a radical Humean who might go so far as to deny even this, claiming that having a conscious life at all is an extrinsic matter. But I've argued that they must surely stop short of accepting that it is *maximally* extrinsic. Given the direct epistemic access we have to our own consciousness, such a view would be completely unmotivated from the Humean perspective and even undermine their own explanatory aims. Therefore, the Humean should at least accept the following, even weaker premise.

P1*. An intrinsic duplicate of *some* region wholly containing me will contain a being with a conscious life.

By opposing the Humean, one need not settle exactly which duplicate region or regions will contain duplicates of me with a conscious life, or even which (if any) regions it would take to duplicate the same conscious life. Once we endorse non-Humean c-n relations, however, the opportunity arises to restrict the relevant regions to those which maintain the same (intrinsic) c-n connections to me. Regardless, P1* is expressed in such a non-committal way as to give the Humean as much room for manoeuvre as they might conceivably allow themselves. As we will see, that's not enough to help them avoid the (revised) argument from consciousness.

4. *Strengthening P2*

We've weakened the first premise of the argument from consciousness to something we can be more certain the Humean will accept. But as it stands, an inference to the falsity of Humeanism from the conjunction of P1* (or P1' for that matter) and P2 to the conclusion is invalid. It is consistent with both

premises that while my conscious life has causal requirements, my duplicate (who may have a qualitatively different conscious life) does not have causal requirements. To compensate, P2 will have to be strengthened somehow to make claims about the causal requirements of my duplicates in relevant duplicate regions whose conscious lives may not be quite like mine.¹² Also, there is unnecessary ambiguity in the phrase (originally in Hawthorne's argument) 'my conscious life': Do we mean my specific conscious experience or rather only my having conscious experiences?¹³ Although I think the defence provided below would work for either interpretation, the latter is all we need. These points considered, I propose the following new second premise.

P2*. There are causal(-nomic) requirements on my having a conscious life and on my conscious duplicates in any duplicate region wholly containing them having a conscious life.

P2* requires acceptance of an instance of the plausible principle that whatever goes for the c-n requirements of some phenomenon goes for its duplicates in intrinsic duplicate regions wholly containing them that also exhibit that phenomenon. At least, the principle implies nothing about the intrinsic–extrinsic status of c-n relations so the Humean can't object to it on that count. Nevertheless, justification for P2* is more demanding than for P2. And it is not obvious that Hawthorne's defences of the latter do a particularly good job of supporting either.

Hawthorne mentions a number of 'a priori obvious' considerations in favour of P2 including those of conscious awareness and the unity of consciousness. Hawthorne takes two of these to be worth expanding on. First, the conscious experience of *attention* to some worldly event is intuitively a causal matter, and the widespread endorsement that there is a necessary causal relation between perceiver and perceived is symptomatic of that intuition. This remains the case regardless of whether or not one is attending to an external event or some internal ongoings, such as a headache. So the fact that I'm now attending to my headache puts a causal requirement on my (current) conscious life.

Second, Hawthorne suggests that an instantaneously existing being could not exist long enough for 'the flame of consciousness to flicker'. But a mere mereological fusion of instantaneous beings would not overcome the inner darkness unless each in the series was causally connected, and the widespread endorsement of causal continuity for personal identity is symptomatic of this intuition. Because the flame of consciousness burns clearly within me, again there must be causal requirements on my conscious life.

Hawthorne also offered a further 'style of consideration', that certain phenomenal states seem to be constituted by causal powers. For instance, 'phenomenal colours are thus disposed to produce certain similarity

verdicts' (p. 355). Supposing causal powers (when acting) give rise to causal relations, then similarity verdicts of this sort require causal relations.

I don't think the Humean will find these considerations very compelling reasons to accept either P2 or P2*. Regarding the first, Humeans might allow that *genuine* attention fails if there is no c-n relationship linking conscious state and thing attended to, while still claiming that this makes no difference to inner conscious life. It may be, for instance, that I am actually attending to my headache because of a genuine causal connection, but it could have been that my conscious experience *as of* a headache and the source of pain in my head were not c-n connected. Moreover, even if the Humean wants to grant (along disjunctivist lines) that the lack of a c-n connection would affect the experience itself, that might be entirely consistent with a denial of P2*, which concerns merely having a conscious life, not any specific experience.

Regarding the second consideration, Humeans tend to be perdurantists: human beings are just a series of instantaneous states, and whether any single one of them can be alone conscious is at least an open question with an empirical answer. In this light, demanding that the states be causally connected for there to be consciousness seems just to beg the question against their position. A similar criticism can be raised for Hawthorne's further style of consideration, because Humeans believe causal powers are reducible to facts about regularities among the ongoing local matters of fact. But if it's the matters of fact which ultimately matter for consciousness, then the Humean will not consider the causal relations established when causal powers manifest to be essential for a conscious life.

So we need another defence for the argument's second premise, specifically, one able to justify P2*. Like Hawthorne's example of instantaneous beings, my defence of P2* rests on a thought experiment. However, my thought experiment is more elaborate than Hawthorne's and also differs in two dialectically important respects. First, instead of making claims about the consciousness of temporal stages, I focus instead on synchronically persisting spatial parts. We may or may not find it plausible that an instantaneous being could be conscious, but we can surely be more confident that a single neuron cannot be (*pace* panpsychists), even more so a single dendrite or sodium channel. Consequently, we should be more confident that a mere fusion of the component parts of appropriate realisers for human consciousness will not necessarily be conscious. But moreover, instead of trying to evoke intuitions about the state of consciousness of fusions of isolated component parts, I focus instead on alternate fusions of beings *stipulated* to be conscious. The crucial question then becomes whether *all* such fusions are conscious. As we will see, this way of setting things up makes the case for P2 and P2* much stronger.

Imagine two conscious human beings *A* and *B* who have brains physically identical in structure, organisation and activity over some period of time Δt

right down to the whatever level is relevant to constituting consciousness. If we like, they can be exact intrinsic duplicates, and we could even stipulate that one of them is me. Let us imagine in particular that all the activity which is constitutive of whatever conscious experiences *A* and *B* have at some time occurs in sections of the cortex that can be divided up (arbitrarily) into two spatially distinct matching pairs X_A and Y_A in *A*'s brain and X_B and Y_B in *B*'s brain (say, the frontal and the parietal lobes). It doesn't really matter which parts of the brain these are, or even that they are part of the brain, so long as we agree that there are *some* parts of *A* and *B* which realise (constitute, give rise to, ground) their respective instances of consciousness but that neither part is individually sufficient for their conscious state. (If you're concerned that the division into two parts is not enough to ensure individual insufficiency, then just run the example with more divisions and more people.)

Given a reasonably liberal mereology, we can also consider the object *C* comprising the mereological sum of X_B and all of *A* except X_A during Δt , and the object *D* comprising the mereological sum of X_A and all of *B* except X_B during Δt . Because *A* and *B* are located at different places, *C* and *D* are, therefore, 'scattered objects', if they exist at all, since they comprise parts that are not all spatially continuous with each other (e.g. X_B is not continuous with any of the rest of *C*).

Although we may be happy to grant their existence, I take it that we should deny that *C* and *D* are conscious. It may be hard to say what would count exactly as evidence either way, but any philosophical view committed to *C* and *D*'s consciousness would be reasonably criticised for resting on wild speculation. In contrast with *A* and *B*, for instance, who can be interacted with, talked to, given Turing-tests, etc., we don't seem to be capable of the same sort of contact with *C* and *D*. It's also hard to see what would constitute them fulfilling the right causal role for any of the mental states we might wish to ascribe to them.

But if it's right – and I will assume it is – that *A* and *B* are conscious whereas *C* and *D* are not, we can ask what makes for the difference. Obviously, *A* and *B*'s brains have a geometry that is in some sense *normal*, both in the arrangement of their neuro-chemistry and their spatial extension; they have the shape and size of a normal brain. Not so for *C* and *D*. Their 'brains' (if it makes sense to refer to them as such) are scattered objects, mediated by non-brain matter, such as the skulls of *A* and *B* and whatever is going on in between *A* and *B*. Given that this is really the only salient distinction, such facts must surely be what motivates the understanding that *A* and *B* but not *C* or *D* are conscious.

What is relevant to the Humean, specifically however, is whether the difference in geometry itself can be what *explains* the difference in consciousness, as opposed to merely signalling that there is a difference. And here I think there's good reason to think that it can't be. After all, each of the

component parts of *C* and *D* have exactly the right configuration of activity to be the components of a conscious life, and they receive the right stimuli from their environment. Why should the fact that there is some particularly large distance between some of the components matter? Things might be different if spacetime was structured in such a way that differences in geometry and proximity amounted to discrete differences between, on the one hand, the relationships instantiated by *A* and *B*'s brains and, on the other hand, *C* and *D*'s. But spacetime is, we suspect, continuous (or at least not discrete at a level relevant to neural chemistry). Hence, *A* and *B* would remain conscious if their brains were a little squished (e.g. under a moderate gravitational field) or otherwise slightly warped (e.g. on a theme-park ride). The synaptic clefts between *A* and *B*'s neurons can also vary by very small amounts without precluding the possibility of consciousness, and we might even expect their conscious lives to survive radical restructuring following from localised brain lesions. So why should it matter that some of the distances involved in *C* and *D*'s 'brains' are a little larger?

Of course, the distances *do* matter. But now I think it's pretty clear that the reason we think *C* and *D* aren't conscious is not due the geometry alone. Rather, it's what the geometry implies about the available c-n relations: the distances involved make certain *interactions* among the components of *C* and *D*'s brains impossible. Folk physics has taught us to be wary of positing, let alone assuming, action at a distance, and because the two components in both *C* and *D*'s brains necessary for consciousness are stipulated to be far apart and (presumably) mediated by all sorts of other goings-on, they simply cannot interact in the right way for consciousness.

That it is really the c-n relations that are at issue can be emphasised by imagining further that some wizard-skilled neuroscientist is able to wire up with super-fast fibre-optic cables the brain-sections X_A and Y_B such that *D*'s brain really *is* connected in at least the same configuration of neural connections as *A*'s and *B*'s are when there aren't the additional wires. In this alternative scenario, I take it we will not be so hasty to deny the consciousness of *D*. Why? because we can no longer be sure that the kinds of causal interactions that happen in actual cases of consciousness can't be happening for *D*. Moreover, presumably we'd also be less likely in this revised situation to think that *A* and *B* are conscious, because the signals necessary to facilitate interaction between X_A and Y_A and between X_B and Y_B have now been tampered with.

Obviously, the thought experiment just provided only supports P2* if intuitions really do align in the way I'm suggesting they will. And intuitions are defeasible. We earlier questioned the sufficiency of Hawthorne's reliance on 'bedrock intuition' in defence of P1, because that intuition was clearly incompatible with numerous views and expressions of intuition voiced in philosophy of perception and consciousness. By contrast, however, the intuition

motivated here does not seem to be. In fact, the intuition seems well supported by both theoretical and empirical scientific practice.

To date, a number of scientific 'models of consciousness' exist, which aim to say what consciousness consists in Seth (2007, 2009). These models are deeply influenced by scientific developments but also aim to make more general claims about what consciousness is independently of specific neural chemistry. We might fairly divide these theories into functionalist and non-functionalist theories, where the former does and the latter does not treat types of mental states as characterised by the type of causal role they play in a larger network linking inputs, outputs and other internal states (Kobes, 2010). For instance, global workspace theory is a prime example of the former and requires one's mental architecture to comprise a central implemented workspace that is able to 'access' other domains. Clearly, any theory that analyses conscious life in terms of the causal role of mental states is committed to P2* and possibly even the more general claim that there are causal requirements on any instance of consciousness (cf. Hawthorne, 2004, p. 354). But non-functionalist theories are liable to be committed to c-n connections for consciousness too. Of these, integrated information theory (IIT) is perhaps the most well-known. It requires, unsurprisingly, a certain degree of *integration* of information across consciousness's constituents. As it is employed by IIT theorists, the notion of integration is purely phenomenal. Nevertheless, it is appreciated that the notion has causal implications for the physical substrate of consciousness (PSC), that is, whatever physical matter implements conscious life. Hence, one of the view's most prominent advocates remarks that 'the postulates of IIT state that the PSC must have intrinsic cause-effect power; its parts must also have cause-effect power within the PSC, and they must specify a cause-effect structure that is specific, unitary and definite' (Tononi *et al.*, 2016, p. 450).

Both global workspace theory and IIT – exemplary candidates on both sides of the functionalist/non-functionalist divide – straightforwardly predict the difference between *A* and *B*, on the one hand, and *C* and *D*, on the other. For only the first two brains comprise fully integrated architectures with a central workspace able to access all the domains instantiated across different constituent parts. But as we have seen, these differences in integration and access mean differences in the c-n relations. This would also explain why a wizard-skilled re-wiring really might exactly switch who has consciousness and who doesn't. So long as we take these models' developers at their word, and the agreement among them as indicative of truth, then the intuition that c-n connections are essential to human consciousness appears to have support beyond mere philosophical intuition.

Empirical support also comes more directly from the neuroscience. Here, practitioners talk frequently of the required *connections* for maintaining consciousness and never, it seems about the importance of geometry *per se*. For

instance, one recent study compared the neural operations of (unconscious) patients suffering from coma-causing brainstem lesions with (conscious) patients suffering non-coma-causing brainstem lesions. They found that injury to a small region in the rostral dorsolateral pontine tegmentum (a part of the brainstem) ‘was significantly associated with coma-causing lesions’ (Fischer *et al.*, 2016, p. 2427). The authors concluded that ‘this brainstem site is functionally connected to two cortical regions, the AI and pACC, which become disconnected in disorders of consciousness’ and that ‘this network of brain regions may have a role in the maintenance of human consciousness’. On the face of it, results like these reflect the idea that it is c-n relations (e.g. between the pontine tegmentum and AI and pACC regions) that are required for consciousness. For instance, the authors of the study explicitly reference ‘disconnection’ between these regions, the failure of a ‘network’, being associated with coma.

In sum, despite it being a difference in the geometrical facts, which signals to us that even if *A* and *B* are consciousness *C* and *D* cannot be, we have a very good reason to think that the reason for this difference is ultimately due to the fact that c-n connections in *A* and *B* are missing in *C* and *D*. This provides strong support for P2. But it also provides strong support for P2*, because it shows that a conscious duplicate of me in an intrinsic duplicate region will also require the c-n connections too. Of course, Hawthorne’s own defences may have anyway persuaded us. But unlike those defences, the case involving *A–D* is the one that makes no contentious claims about the nature of attention, instantaneous beings or causal powers. It is also able more straightforwardly to draw on the support of empirical and theoretical developments in the science of consciousness.

5. *The consistency of the premises and a lesson for the non-Humean*

With P1* and P2* both argued for independently, it’s worth asking whether there might be any issues of compatibility. One potential issue is that P1* was argued for by drawing attention to differences in the observability of consciousness and c-n relations while P2* argued that consciousness is in fact dependent on c-n relations.¹⁴ More carefully, if P2* is right, then the c-n relations between constituent parts of me are essential to my being conscious, and of course it is my being conscious that allows me to tell that I’m conscious. But if the argument for P1* is right, then what allows me to tell that I’m conscious is something that marks an observable difference from whatever the world would have to be like for me not to be conscious. In combination, this means that c-n relations must make for an observable difference, because they are instrumental in giving rise to something

(consciousness) that makes for observable differences. However, in arguing for the observable difference afforded by consciousness, I granted that both Humeans and non-Humeans are inclined to accept that two regions duplicated up to the inclusion and distribution of c-n relations would be observationally alike. If they're right to do so, we face contradiction.

Thankfully – although not for the Humean! – the contradiction only comes about by granting things that need not be granted in order for the argument against the Humean to go through. As just remarked on, in arguing for P1* it was pointed out that Humeans and non-Humeans alike tend to agree that c-n relations don't make for observable differences. But this was really just a concession to the Humean in order to show that an independently plausible view about c-n relations doesn't extend to consciousness. Granting the lack of observable differences between regions differing only in how c-n relations are distributed among them, the Humean has a good explanation for why that is, namely, c-n relations' maximal extrinsicness. This sort of reasoning does not carry over to the case of consciousness, however, because my conscious life clearly does make for observable differences. This all puts the Humean in trouble because they are *committed* to the maximal extrinsicness of c-n relations (it is a defining characteristic of their view) and yet, according to P2*, consciousness is dependent on them.

The non-Humean is not in the same kind of trouble. Admittedly, they would be in trouble if they wanted to grant that c-n relations don't make for observable differences. For regardless of whether or not one grants c-n relations' maximal extrinsicness it seems inconsistent to say that something that makes for no observable differences grounds something that does. But the non-Humean need not after all grant the assumption. Rather, the non-Humean should say that *some* c-n relations make for observable differences, specifically, the ones involved in relating the constituents of my consciousness. Perhaps this is not something non-Humeans have typically noted, but it is rendered highly defensible by the foregoing discussion.

The claim that c-n relations can make for observable differences is, moreover, not liable to a wildly implausible epistemology. Non-Humeans need not, for instance, accept an observable 'tie' or 'spark' between causes and effects. Nor need non-Humeans find the defeasibility of causal-nomic conjectures in science strange. For instance, when c-n relations are not those involved in the constituents of my very own consciousness, then there is no reason to think they would make for any observational differences, for me at least. Moreover, were I to look via magnetic resonance images (or whatever) into my own neurochemistry, there is no reason to suppose that I should be able to tell via observation that the observed causal relata are those making for the observational differences afforded by my awareness of my own consciousness. This is because it is not directly perceivable which neurochemical phenomena are in fact constitutive of my concurrent consciousness episodes. So the non-Humean can be committed to at least some

c-n relations having the kind of tell-tale observational mark which Humeans can't allow without having to come into conflict with commonly accepted epistemological limitations of scientific discovery.

In sum, the arguments provided for the revised premises are consistent so long as we accept that some c-n relations make for observable differences. That is something the Humean cannot accept. But it is something the non-Humean can and should.

6. *Conclusion*

With the revised premises, the case against Humeanism now goes as follows:

- P1*. An intrinsic duplicate of any region wholly containing me will contain a being with a conscious life.
- P2*. There are causal(-nomic) requirements on my having a conscious life and on my conscious duplicates in any duplicate region wholly containing them having a conscious life.

Therefore, Humeanism is false.

The conclusion follows for much the reasons Hawthorne originally noted. A duplicate of me in some duplicate region must be conscious. But because that duplicate's consciousness will be dependent on c-n relations within the duplicated region, it had better be that c-n relations are preserved in all cases of duplication. But if GENERALISM is right, then (*pace Weatherston*) there could be facts elsewhere in the world which undermine the regularities required for those c-n relations to hold. For instance, the duplicated region could be of the entire actual world located within Plenitude world, where the laws are the theorems of PLENITUDE. The laws in such a world are just not sufficient to provide the c-n relations present in our world. Unless the Humean is therefore willing to countenance singularism, their position is untenable.

What of the non-Humean? Their position is stable so long as they are not drawn into conceding to the Humean that c-n relations never make for observational differences. If P2* is true, *some* c-n relations must make for observational differences, because my consciousness makes for such differences. That is a lesson for the non-Humean which I've suggested is entirely plausible, if surprising.

7. *Acknowledgements*

I would like to thank Francesca Bellazzi, Alexander Franklin, Max Jones, Samuel Kimpton-Nye, Vanessa Seifert, Neil McDonnell, Tuomas Tahko,

Leila Tufekci and an anonymous reviewer for enormous help on my development of the ideas in this article. This research has received funding from the European Research Council under the European Union's Horizon 2020 research and innovation programme, grant agreement 771509.

Department of Philosophy
University of Bristol

NOTES

¹ I suppose the argument *could* be formulated in more general terms: there are *some* human agents whose conscious lives are (1) preserved under duplication and (2) have causal requirements. This might be more dialectically appealing in some respects but can also suggest there is a significant question over why those agents satisfy both characteristics rather than only one, which would distract from the issue at hand. Satisfying both characteristics is an entirely commonplace feature of ordinary consciousness.

² Citations are slim, although articles invoking but not directly responding to his argument have been recently in print, including Shumener (2019) and Pallies (2019).

³ Whether or not worlds should be allowed to be infinitely compounded by recombinations is a point of contention; refer to Lewis (1986b) and Russell and Hawthorne (2018).

⁴ Pallies draws attention (note 11, p. 147) to the concern that Lewis's own views on recombination would have rendered such a world impossible. Russell and Hawthorne (2018), however, defend its possibility. Moreover, it is surely plausible that more 'modestly plenitudinous' worlds can be described in which the same point goes through.

⁵ Arguably, the notion of 'trumping' pre-emption is otiose, because such scenarios can be conceived as a form of the more mundane early or late forms of pre-emption (refer to Bernstein, 2015).

⁶ Intrinsic and extrinsic are exclusive and (I will assume) exhaustive. Defining them is a topic in itself, but for our purposes, we can keep in mind the rough and ready idea that a property or relation is intrinsic to its bearers if it is had independently of whatever may or may not accompany its bearers (Lewis and Langton, 1998).

⁷ I certainly don't want to presume that the non-Humean *must* offer such a poor explanation. The non-Humean might simply refuse to countenance regions differing *only* in the c-n content. Alternatively, they might accept that there are observational differences, contrary to the assumption here. Refer to Section 5.

⁸ The restriction to an actual conscious life is important. We no longer believe in ectoplasm or vital forces. Consequently, we think that duplicate regions in which actual conscious humans are duplicated will not include such things. That is why (I argue) good old-fashioned c-n relations must actually do the work of underlying consciousness even if there could be counterparts of actual conscious humans in regions where these alien properties or substances do play the supporting role.

⁹ Lewis also seemed to admit the intuition Hawthorne voices and made his peace with his denial of it: 'if laws underlie causation, that means that we are wrong if we think, for instance, that the causal roles of my brains states here and now are an entirely local matter. That's an unpleasant surprise, but I'm prepared to bite the bullet' (Lewis, 1994, p. 479).

¹⁰ Hawthorne considers an argument with this first premise and a second premise similar to that I offer below (refer to his footnote 9). He rightly points out that the first becomes easier to defend while the latter becomes harder. The new defence of the second (revised) premise in Section 4 aims to take up the challenge of this harder defence.

¹¹ The key point can even be adopted by so-called ‘eliminativist materialists’ who fear that the word ‘consciousness’ is too muddled to be useful in future science (Churchland, 1983; Wilkes, 1984). If Humeans take this stance, then they should at least accept that there is some accessible difference between being such that the folk would call me ‘conscious’ and being such that they would not. All that is required of what I am referring to as consciousness is that it is something about humans’ mental lives which they can be aware in a rather direct way that they ‘have’, in some sense, and that humans with brains not functioning in quite the right way may not end up having in the same sense. Consciousness in this sense is just whatever neuroscientists are looking for when they look for the ‘correlates of consciousness’.

¹² It would be sufficient for the revised second premise to assert that there are causal requirements on consciousness in general, but that is more than we need and arguably false (refer to footnote 8 above, and Hawthorne’s footnote 9).

¹³ Thanks to an anonymous referee for suggesting the disambiguation. In context, the latter interpretation can be read as making claims about the c-n grounds for my being conscious.

¹⁴ Thanks to an anonymous reviewer for drawing this potential issue to my attention.

REFERENCES

- Anscombe, G. E. M. (1971). *Causality and Determinism*. Cambridge: Cambridge University Press.
- Bernstein, S. (2015). ‘A Closer Look at Trumping,’ *Acta Analytica* 30(1), pp. 1–22.
- Braddon-Mitchell, D. (2001). ‘Lossy Laws,’ *Noûs* 35(2), pp. 260–277.
- Churchland, P.S. (1983). ‘Consciousness: The Transmutation of a Concept,’ *Philosophical Quarterly* 64, pp. 80–95.
- Cohen, J. and Callender, C. (2009). ‘A Better Best Systems Account of Lawhood,’ *Philosophical Studies* 145, pp. 1–34.
- Dretske, F. (1995). *Naturalizing the Mind*. Cambridge, MA: MIT Press.
- Ducasse, C. (1926). ‘On the Nature and the Observability of the Causal Relation,’ *Journal of Philosophy* 23(3), pp. 57–68.
- Fischer, D., Boes, A., Demertzi, A., Evrard, H., Laureys, S., Edlow, B., Liu, H., Saper, C., Pascual-Leone, A., Fox, M. and Geerling, J. (2016). ‘A Human Brain Network Derived from Coma-Causing Brainstem Lesions,’ *Neurology* 87, pp. 2427–2434.
- Frankish, K. (2016). ‘Illusionism as a Theory of Consciousness,’ *Journal of Consciousness Studies* 23(11–12), pp. 11–39.
- Hawthorne, J. (2004). ‘Why Humeans Are Out of their Minds,’ *Noûs* 38(2), pp. 351–358.
- Kobes, B. (2010). ‘Functionalist Theories of Consciousness,’ in T. Bayne, A. Cleeremans and P. Wilken (eds) *Oxford Companion to Consciousness*. Oxford: Oxford University Press.
- Lewis, D. (1973). *Counterfactuals*. Oxford: Blackwell Publishing.
- Lewis, D. (1983). ‘New Work for a Theory of Universals,’ *Australian Journal of Philosophy* 61 (4), pp. 343–377.
- Lewis, D. (1986a). ‘Mad Pain, and Martian Pain,’ in *Philosophical Papers I*. New York: Oxford University Press, pp. 122–132.
- Lewis, D. (1986b). *On the Plurality of Worlds*. Oxford: Blackwell.
- Lewis, D. (1994). ‘Humean Supervenience Debugged,’ *Mind* 103, pp. 473–490.
- Lewis, D. and Langton, R. (1998). ‘Defining ‘Intrinsic’,’ *Philosophy and Phenomenological Research* 18(2), pp. 333–345.

- Loewer, B. (1996). 'Humean Supervenience,' *Philosophical Topics* 24, pp. 101–127.
- Loewer, B. (2007). 'Counterfactuals and the Second Law,' in H. Price and R. Corry (eds) *Causation, Physics, and the Constitution of Reality: Russell's Republic Revisited*. Oxford: Oxford University Press, pp. 293–326.
- Menzies, P. (1999). 'Intrinsic Versus Extrinsic Conceptions of Causation,' in H. Sankey (ed.) *Causation and Laws of Nature*. Dordrecht: Kluwer Academic Publishers, pp. 313–330.
- Pallies, D. (2019). 'Why Humean Causation Is Extrinsic,' *Thought* 8, pp. 139–148.
- Price, H. and Corry, R. (eds) (2007). *Causation, Physics and the Constitution of Reality: Russell's Republic Revisited*, Oxford: Oxford University Press.
- Russell, J.S. and Hawthorne, J. (2018). 'Possible Patterns,' in K. Bennett and D. Zimmerman (eds) *Oxford Studies in Metaphysics: Volume*, Vol. 11. Oxford: Oxford University Press.
- Schaffer, J. (2000). 'Trumping Pre-emption,' *Journal of Philosophy* 98, pp. 75–92.
- Seth, A. (2007). 'Models of Consciousness,' *Scholarpedia* 2(1), p. 1328. revision #132493.
- Seth, A. (2009). 'Explanatory Correlates of Consciousness: Theoretical and Computational Challenges,' *Cognitive Computation* 1, pp. 50–63.
- Shumener, E. (2019). 'Humeans Are out of this World,' *Synthese*, <https://doi.org/10.1007/s11229-019-02439-8>
- Steward, H. (2011). 'Perception and the Ontology of Causation,' in J. Roessler, H. Lerman and N. Eilan (eds) *Perception, Causation and Objectivity*. Oxford: Oxford University Press.
- Tononi, G., Boly, M., Massimini, M. and Koch, C. (2016). 'Integrated Information Theory: From Consciousness to its Physical Substrate,' *Nature Reviews Neuroscience* 17(7), pp. 450–461.
- Weatherston, B. (2007). 'Humeans Aren't out of their Minds,' *Noûs* 41(3), pp. 529–535.
- Wilkes, K.V. (1984). 'Is Consciousness Important?,' *British Journal of Philosophy of Science* 35, pp. 223–43.
- Wilson, J. (2009). 'Resemblance-Based Resources for Reductive Singularism,' *The Monist* 92(1), pp. 153–190.