

Patrick Lewtas

# *Passive Causation*

## *Making Interactionism Work*

**Abstract:** *This paper advances a theory of interactionist mental causation within a non-physicalist property dualist framework. It builds from Chalmers' argument that non-physical experiences can't have causal powers. It reinterprets this as a constraint, then identifies the unique model satisfying it. This has the experience existing passively with physical nature responding actively to it. The paper explores the causal theory presupposed thereby; applies its model to standard property dualism (emergentism) and particulate property dualism (panpsychism); shows how the model coheres with science; integrates it into a revised functionalist framework having three levels instead of the usual two; explains why consciousness detectors will remain impossible; and argues that dualists should find the model troubling.*

### **1. Problems of Mental Causation**

Most believe conscious mental entities make things happen. Your thirst leads you to drink. Napoleon's vanity littered Europe with death. But none sees clearly how.

The physicalist forges the problem into the causal argument. Because (1) conscious properties exist and cause physical events; (2) all physical events (understood as physical properties instanced at

Correspondence:  
Email: [plewtas@gmail.com](mailto:plewtas@gmail.com)

*Journal of Consciousness Studies*, 25, No. 9–10, 2018, pp. 139–62

times<sup>1</sup>) that have causes have sufficient physical causes (the doctrine of physical causal closure); and (3) systematic overdetermination doesn't or can't occur;<sup>2</sup> the physicalist identifies (token) conscious properties with (token) physical properties. Either conscious thoughts have a covert physical nature or they do nothing.

But perhaps physicalism entails the denial of mental causation. Davidson, for instance, allows that conscious thoughts, as physical entities, do cause, but leaves us wondering whether they cause *as conscious thoughts*.<sup>3</sup> Kim — who defends physicalism to safeguard mental causation — tells us that conscious thoughts and their physical effects are higher-level supervenient entities, while all causation happens at the bottom-most micro-physical level (see Kim, 2000, chapters 2 and 3). The conscious thought apparently does nothing.

Do non-physicalist dualists fare better? They must impugn the causal argument's validity or deny one of its premises. Some forswear mental causation (Jackson, 1982; Chalmers, 1996; Kim, 2005). But they pay a steep price for theory: powerless minds fly in the face of evolution and common sense, leave opaque how we know ourselves, and doom agency. This paper judges this prohibitive. Those hoping to retain interactionist causation have three options. (1) Abandon physical causal closure (e.g. Lowe, 2013). (Some then explain how mental causation could cohere with the evidence for closure.<sup>4</sup>) (2) Reject non-overdetermination — by arguing that, rightly understood, overdetermination does allow for interactionist causation (Mills, 1996). (3) Or contend that the causal argument's premises don't entail its conclusion.<sup>5</sup> The path then lies open to a property dualism where

- 
- <sup>1</sup> We could recast the argument to accommodate other understandings of events or reformulate it in terms of substance-causation rather than event-causation. These changes wouldn't matter here.
  - <sup>2</sup> Systematic overdetermination occurs when each token of every mental cause is overdetermined by some physical cause — and not necessarily the same kind each time.
  - <sup>3</sup> See Kim (1995; 2010), as well as the defences in Davidson (1995) and Gibb (2013a).
  - <sup>4</sup> Averill and Keating (1981) venture that mental causes comply with conservation laws by leaving the quantity of energy untouched while changing its distribution. Larmer (1987) exempts mental causation from conservation laws on the grounds that human bodies don't constitute the kinds of closed systems to which the laws apply. Eccles (1987) suggests that mental causes skew quantum probabilities while keeping outcomes within the chances allowed by physical law. Stapp (1993) adverts to consciousness's alleged role in the collapse of the wave function.
  - <sup>5</sup> Lowe (2008) (a substance dualist) denies the ontological homogeneity of the causal relata presupposed by the causal argument. Crane (1995) (a non-reductive physicalist) denies the ontological homogeneity of the causal relation presupposed by the causal

physical entities instantiate non-physical conscious properties. Perhaps these conscious properties emerge to characterize suitably organized brains. Or maybe they characterize basic physical entities — like electrons and quarks — then come together to build complex minds when the basic physical entities come together to build brains.

But wait! The dualist still faces a hurdle. Call this *Chalmers' problem* (Chalmers, 1996, pp. 157–8).<sup>6</sup> It purportedly establishes the impossibility of mind–body interaction on the grounds that non-physical conscious thoughts can't have causal powers. The first part of Chalmers' problem speaks to the warrant for non-physicalism. The dualist can treat consciousness as non-physical only if he allows it a non-physical nature over and above any causal/functional properties it might (seem to) have. After all, if he identifies consciousness with some set of causal/functional properties, then consciousness amounts to a role needing something to realize it. But nothing stops the physical from doing this.<sup>7</sup> So the conscious thought ends up (at least token) physical. Or at least the dualist loses any reason for thinking otherwise. Further, what holds of the overall conscious thought holds of its parts: any part that doesn't stand over and above the causal/functional yields to (the possibility of) physical realization and therefore (at least token) physical reduction. In sum, to the extent the dualist identifies consciousness with causal powers, to that extent he undercuts his motivation and warrant for rejecting physicalism.

The second part of Chalmers' problem speaks to the metaphysics of dualism. The first part forces the dualist to give consciousness a non-causal/non-functional nature — something he would do anyway. But if the dualist also accords consciousness a causal power, then he must explain what relation its non-causal/non-functional essence bears to its essentially causal/functional power. He can't identify these on pain of incoherence, since nothing essentially non-causal/non-functional can equal anything essentially causal/functional. He must instead

---

argument. Gibb (2013b) construes mental causation as a kind of double prevention — the mental cause prevents another mental cause from preventing the actual physical cause — consistent with the causal argument's premises but not its conclusion.

<sup>6</sup> Lewis (1980/1990, p. 513) anticipates Chalmers.

<sup>7</sup> Unless the consciousness-role outstrips the wherewithal of the physical. Then the physical would realize many roles, but something else — neither conscious nor physical — would realize consciousness. Another possibility: something non-physical realizes the role even though the physical could. The present paper ignores these wrinkles because Chalmers' problem applies equally to them. We still have a conscious property and a distinct causal power contingently related thereto.

recognize them as two separate entities, with disparate natures, only contingently wrapped together.

This has big consequences. Take a conscious property and any of its alleged causal powers. Suppose the property of willing a spoon to bend and the power of bending it through non-physical mind-power alone. At first we think of this as a single entity — a non-physical conscious property essentially tied to the matching non-physical causal power. But Chalmers' problem shows that we really have two essentially distinct properties here. So we can't view the conscious property as the power of bending spoons, or as metaphysically necessitating the power, or as linked so tightly to it that the two can't sunder. Chalmers' problem seems to prove that a conscious property can't have causal efficacy in its own right — that dualism entails epiphenomenalism.<sup>8</sup>

Chalmers' problem doesn't arise for physical properties because their causal/functional natures exhaust them — at least as science knows them. As science has it, high-level physical properties — being a liquid, say — consist of high-level causal/functional roles built of (or realized by) lowest-level physical properties themselves consisting of causal/functional roles. Thus science understands the electron only as a bundle of spatio-temporally located dispositions to affect other things thus-and-so and to be affected thus-and-so by other things. Science can't say more. It can only gather data about manifested dispositions — because all data must causally affect scientists' senses (if not directly, then indirectly through prior effects on instruments)<sup>9</sup> — and then explain them by postulating entities disposed to cause them. Hypotheses framed in terms of non-dispositional entities would lack explanatory power and stand aloof from empirical test. Science thus holds silent about any intrinsic or non-dispositional properties that might fill out physical structure, ground physical dispositions, or characterize physical nature apart from its relations to other things. (Maybe such entities exist. But science can't access them, recognize

---

<sup>8</sup> Chalmers' problem doesn't concern the possibility that an effect of consciousness might have been caused by something else, perhaps something physical. Nothing requires that an effect produced by consciousness be producible only by consciousness. Chalmers' problem instead highlights the fact that causal powers, allegedly of consciousness, can't relate essentially to the 'conscious part' of consciousness.

<sup>9</sup> Might manifested dispositions have non-dispositional properties where dispositions don't? It doesn't matter. Even if non-dispositional manifestations occur, science can know nothing of them. The scientist can only learn about entities disposed to affect the senses.

them, provide evidence for them, or tell us anything about them.) Thus physical reality, as portrayed by science, consists solely of spatio-temporally located dispositions and arrangements thereof.

Now, if physical nature doesn't reach beyond the causal/functional, then of course the physical can't peel away from the causal/functional. Chalmers' problem arises only in the case of consciousness, where the dualist insists that conscious nature does outrun the causal/functional.

Have we hit a wall? Physicalism purports to save mental causation, but drops the mental from the causation. Dualism does worse, rendering mental causation impossible. Do conscious entities really do nothing?

## 2. Motivations for Passive Causation

Chalmers regards (what we call) Chalmers' problem as disproof of dualist interactionism. This paper reinterprets the problem as a constraint. Chalmers would say the constraint narrows the possibilities to none. It certainly rules out active conscious powers, like a fifth fundamental force unique to consciousness. This paper shows, however, that Chalmers' problem affords a loophole.

The case of the physical points the way. If Chalmers' problem doesn't arise there because physical nature can't separate from the causal/functional, then interactionism will work only with mental powers inseparable from conscious nature. Chalmers' problem proves the interactionist can't resort to anything consciousness does — because any such doings fall within the causal/functional. That leaves only what consciousness *is*. But most dualists insist that each conscious experience has a nature, exists in some determinate way, has content — that it could differ and therefore carries information. Here lies the needed *is*. Interactionism requires this nature, this content, to make a difference — without actually doing anything.

This can work only if consciousness has *passive* efficacy.<sup>10</sup> It sits there as is, and the physical responds to it. The experience doesn't exert force upon the physical the way a flying brick exerts force on a

---

<sup>10</sup> The present paper's active/passive distinction doesn't match the homonymic distinctions in the causal powers literature. For example, Lowe (2013, pp. 158–9) treats as active any power that doesn't need triggering circumstances in order to manifest (a radium atom's power to decay) and as passive any power that does need triggering circumstances in order to manifest (salt's power to dissolve). The present paper distinguishes, instead, between powers that 'exert themselves' — whether prompted or not — and powers that don't. It thus classifies both of Lowe's powers as active.

windowpane. Instead, the physical ‘detects’, ‘reads’, or otherwise ‘picks up on’ the nature/content of the experience and changes its own state accordingly. This keeps the actively causal/functional wholly on the physical side.

### 3. Metaphysics of Passive Causation

What is passive efficacy? What sets it apart from the active efficacy of everyday physical causation? We need some background to answer these questions.<sup>11</sup>

Theories of causation must take one of three mutually incompatible forms. They can trace the source of causation to the causally-interacting entities. They can trace the source of causation to something else. Or they can deny that causation exists. These three exhaust the logical possibilities.<sup>12</sup> The present paper calls the first kind ‘strong’ accounts and the second and third ‘weak accounts’ for reasons soon to become clear.

Consider first those weak accounts that deny causation. These don’t allow that anything ever makes anything happen. The concrete world consists of wholly local and inert entities spread through space and time. These entities together form a contingent pattern — contingent because they might have stood in any other pattern. This pattern has no metaphysical standing in its own right; it exists only because it supervenes on the local entities arranged as so. As it happens, the pattern that actually obtains has features that let us describe (at least parts of) it using dependency relations like regular succession, counterfactual dependency, and/or conditions necessary and sufficient. Because nothing makes anything happen, however, these dependency relations mirror correlations within the pattern, nothing more. Further, if nothing brings anything about, then nothing brings about the overall pattern. The pattern therefore obtains through chance. These weak accounts thus reduce causation to — or eliminate it in favour of — a static pattern that issues from luck. Most philosophers call these accounts ‘neo-Humean’.<sup>13</sup>

---

<sup>11</sup> The paper adapts the next paragraphs from Lewtas (2017).

<sup>12</sup> Transference, probabilistic, and primitivist accounts therefore fall within our three groups, depending on their specific elaborations.

<sup>13</sup> Neo-Humean works abound. See, for example, Lewis (1973), Mackie (1980), Prior, Pargetter and Jackson (1982), and Psillos (2002). Ellis (2001), Mumford (2004), and Strawson (2008) give good overviews of the neo-Humean account.

Weak accounts of the second kind grant that causation exists, but trace its source to something other than the causally-interacting concrete entities. It doesn't make sense, then, to trace causation to other concrete entities. You wouldn't end up with a believable account, for instance, if, having denied that the brick shatters the window, you claimed that Andromeda Galaxy shatters the window when the brick strikes it. Or at least, to save your account, you'd have to allow that Andromeda somehow figured among the causally-interacting entities. These weak accounts therefore locate the source of causation outside the concrete world. This forces them to understand the concrete world as neo-Humeans do: as an arrangement of strictly local entities without powers or intrinsic or essential causal connections to anything. The non-concrete causal source then imposes causal relations on these otherwise inert concrete entities. We can label this source the 'laws of nature', bearing in mind that we don't thereby say much about it. Note that these laws must hold contingently. Since they alone have the wherewithal 'to push things around', nothing exists that could bring them into existence — cause them — or shepherd them into any pattern. It follows that any other laws, even none at all, could have held instead. It follows further that the actual laws (necessarily!) exist by chance. Now, the pattern of laws in our world has features that let us describe the resulting causal connections in terms of dependency relations between cause and effect. Causation thus amounts to chance patterns which we read as dependencies. We see that not a lot sets apart these accounts from neo-Humean accounts, even though metaphysically robust laws do mediate between chance and concrete happenings. We therefore call them 'quasi-Humean'.<sup>14</sup>

Strong accounts make one central claim, *powers*, which holds that robust causes issue from the intrinsic and essentially empowered natures of causally-interacting concrete entities.<sup>15</sup> These empowered

---

14 Today's favoured quasi-Humean accounts understand concrete entities' monadic properties as instances of first-order universals. They then postulate second-order universals that relate the first-order universals. These serve as the laws of nature supplying the concrete world with causal glue. See Dretske (1977), Tooley (1987), and Armstrong (1983).

15 This leaves open whether these natures also have non-causal qualitative properties. Shoemaker (1980) and Mumford (2004) argue that all properties exist only as powers. Williams (2011) contends that no properties qualify as powers. Heil (2003) and Strawson (2016) hold that every property takes the form of a powerful quality — a quality numerically identical to a power. Ellis (2001) and Molnar (2003) allow for non-qualitative powers and non-powerful qualities. Prior, Pargetter and Jackson (1982) insist

natures either ‘reach out’ into the world or ‘stand open’ to the reachings of other such natures. Causation occurs when they engage and interact with one another.<sup>16</sup> Do strong accounts make causation a necessary relation? Yes and no. Empowered natures necessarily reach out and/or stand open, and thus essentially tend to connect with other such natures. But whether they do connect, and with what, and to what effect, turns on what else reaches out and stands open. This remains contingent.

The present paper now shows that passive causation can work only if a strong causal account holds true. Consider neo-Humean accounts first. These do away with causation but reduce its appearance to (our take on) the dependency relations with which, through good luck, we can characterize the spatio-temporal layout of the world’s concrete existences. Because this layout holds contingently — any other pattern might have held instead — the same concrete entities might have worn different causal profiles. Put otherwise, the same concrete entities might have been governed by different natural laws. This yields two points. First, causation — in so far as we consider it real — exists external to the causally-interacting entities. Second, causation rests on formal dependency relations which alone mark off something as cause or effect. These two points help us see that passive causation conflicts with neo-Humean accounts. For these accounts rule out even the possibility of distinguishing passive from active causes, and therefore even the possibility of recognizing passive causation as a distinct kind. This follows from the fact that the so-called passive cause bears the same formal relation to its effect as does any active cause. And since, on neo-Humean accounts, this relation alone constitutes something a cause, the ‘passive’ cause counts as a cause in the same way as an ‘active’ cause.

Neo-Humean accounts can’t accommodate passive causation because they anchor causation in something essentially independent of the cause and the effect. And yet a passive cause differs from an active cause by doing things differently, by being a different kind of causal relatum. The defender of passive causation therefore needs a causal account that grounds causation in the causal relata themselves: in the

---

that qualities and powers never overlap. The interactionist model explored here requires that at least some qualities lack powers — otherwise conscious properties couldn’t have merely passive efficacy. Beyond that the paper remains agnostic.

<sup>16</sup> Bird (2007), Ellis (2001; 2002), Martin (2008), Molnar (2003), Mumford (1998; 2004; 2009), and Mumford and Anjum (2010) present such accounts.



cause and the entity-suffering-the-effect, in their natures and properties. This can work only if the nature of the cause and the nature of the entity-suffering-the-effect interact so as to make the effect happen. Of course, if these things do determine the causal relation between them, then the cause does necessitate its effect, albeit only under the applicable circumstances and in accordance with the relevant natures. Not only does this leave no room for laws of nature that float free of the world's concrete entities while setting their causal profiles. But more, it guarantees the truth of *powers*, the centrepiece sufficient for strong causal accounts. It therefore seems that passive causation can occur only if some strong account holds true.

What about quasi-Humean accounts? If passive causation entails a strong account, and if, as noted earlier, strong accounts and quasi-Humean accounts contradict one another, then quasi-Humean accounts rule out the possibility of passive causation. But let's show this directly.

Quasi-Humean accounts, like their neo-Humean counterparts, ground causation in entities external to the causally-interacting entities. As they have it, non-concrete entities inject causal relations into the concrete world. This happens when a non-concrete causal source links concrete entities of one kind to concrete entities of another kind such that — in the simplest case — the obtaining of an entity of the first kind necessitates the obtaining of an entity of the second. The first entity thereby becomes a cause and the second its effect. As we saw, quasi-Humean accounts (necessarily) hold that the non-concrete entities exist contingently. Quasi-Humeans therefore agree with neo-Humeans that the same concrete entities, with the same intrinsic natures, would wear different causal profiles under different nomological regimes. But now we have the same two points upon which the argument turned in the neo-Humean case. First, causation exists external to the causally-interacting entities. And second, an entity qualifies as a cause or effect only because something else marks it off as such. We now see that quasi-Humean accounts rule out even the possibility of distinguishing passive from active causes. For the so-called passive cause bears the same relation to the relevant non-concrete entity as does any active cause — whatever relation it is whereby the non-concrete entity links cause to effect. On quasi-Humean accounts, then, the 'passive' cause qualifies as a cause in the same way as an 'active' cause.

The same diagnosis applies. Quasi-Humean accounts can't accommodate passive causation because they anchor causation in

something essentially independent of the cause and the entity-suffering-the-effect. And yet a passive cause differs from an active cause by doing things differently, by being a different kind of causal relatum. These same premises lead through the same arguments to the same conclusions. Passive causation entails a strong causal account and rules out a quasi-Humean account. Henceforth the paper assumes the truth of strong accounts.

Now, in both passive and active causation, the entity-suffering-the-effect responds to something-or-other by manifesting a disposition to do such-and-such under the circumstances responded to. Any difference between passive and active causation must therefore concern the something-or-other to which the entity-suffering-the-effect responds. We can see how this must go by comparing what happens in everyday active causation to what must happen in passive causation. Take a simple example of everyday causation — an ice cube melting in a tumbler of warm water. We might say the warm water (the cause) melts the ice (the entity-suffering-the-effect). But we could just as well say the ice (the cause) cools the water (the entity-suffering-the-effect). As others have noted,<sup>17</sup> we in fact have entities (water, ice, air, tumbler, energy, etc.) that interact with one another courtesy of their causal powers — their dispositions to do or to become something in circumstances thus-and-so. So instead of a well-defined cause and a well-defined effect, we have many entities each of which both causes and suffers effects.<sup>18</sup> Everyday active causation therefore features complex symphonies of entities evolving into other complex symphonies courtesy of the relevant causal powers.

Passive causation can't work this way because the cause, the conscious property, can't do anything. It just exists as is. This shunts all activity onto the entity-suffering-the-effect. That entity must 'notice', 'detect', or 'register' the passive cause, then change itself in response. So here, and only here, we have a sharp metaphysical divide between cause and (entity-suffering-the) effect. Note also that passive causation features pairs of interacting entities in place of the ill-bounded (or unbounded) symphonies of active causation. Of course, any given passive cause might influence many physical entities. But each pairing exists as a discrete interaction only contingently related to the others.

---

<sup>17</sup> Mumford (1998; 2009), Molnar (2003), and Martin (2008).

<sup>18</sup> The law of conservation of mass-energy and Newton's third law both entail that a physical entity affecting another physical entity undergoes change itself.

This yields the following picture. A cause causes actively when the relevant causal relations issue, partly at least, from its nature (properties), and a cause causes passively when the relevant causal relation issues entirely from the nature (properties) of the entity-suffering-the-effect. So, with active causation, the properties of the cause give rise to — or contribute to the establishment of — a relation between the cause and the other interacting entities, and this relation triggers various dispositions in the interacting entities (including what we think of as the entity-suffering-the-effect). With passive causation, on the other hand, the properties of the entity-suffering-the-effect give rise to a relation between that entity and the cause, and this relation then triggers some disposition in the entity-suffering-the-effect. So the entity-suffering-the-effect does two things: it actively establishes a relation between itself and the cause, and it actively responds to that relation. The cause passively supplies the opportunity — the occasion, the pretext, the necessary background circumstances — which the entity-suffering-the-effect needs before it can establish a causal relation with it.

#### 4. Mechanics of Mind–Body Interaction

The dualist must overcome at least two hurdles to make this account work. First, he must square it with our best science. His story must cohere with the various conservation laws, especially those of mass-energy and momentum.<sup>19</sup> Second, he must respect the privacy of mind, the fact that — as best we know — only a subject has direct acquaintance with his own consciousness. It won't do to postulate a kind of mind-reading without analogue in daily life. It certainly won't do to postulate mind-reading at a distance.

We most easily descry the outlines of an acceptable account by working through some examples. We attend first to the version of dualism according basic physical entities simple experiences which build complex experiences when the basic physical entities build brains. Consider the bottom-most level. Imagine some basic physical particle A in some quantum superposition of purely physical states  $S_1$  and  $S_2$ . Now imagine that A comes to have experience  $E_1$  when it

---

<sup>19</sup> He must do better than Descartes, who had non-material mind deflecting the paths of physical particles inside the pineal gland — a process he called 'swerve' — in violation of the conservation of momentum.

could have had  $E_2$  instead.<sup>20</sup> Suppose that A's physical nature responds to  $E_1$  by assuming physical state  $S_1$ . Had A instead experienced  $E_2$ , its physical nature would have responded by assuming  $S_2$ .<sup>21</sup> Note that  $E_1$  doesn't do anything active. It doesn't transfer any physical quantity. Its nature doesn't 'reach out'. The physical change — from superposition of states  $S_1$  and  $S_2$  to determinate state  $S_1$  — occurs because A's physical nature first detects and then responds to  $E_1$ . So A's physical nature does all the work while  $E_1$  serves merely as pretext or trigger. This example clears the hurdles mentioned earlier. It doesn't violate conservation laws, since no conserved quantity is created, destroyed, or transferred. And it respects the privacy of consciousness. A's physical nature responds to A's experience. Everything happens internal to A.<sup>22</sup>

Consider another example. Suppose quantum-entangled but otherwise basic physical particles A and B stand at some distance from one another. The two exist in a quantum superposition of purely physical states  $S_1$  and  $S_2$  such that one particle will assume one state as soon as the other particle assumes the other state. Now suppose particle A comes to have experience  $E_1$  when it could have had  $E_2$  instead. And suppose, as before, that  $E_1$  and  $S_1$  relate lawfully, as do  $E_2$  and  $S_2$ . Particle A thus responds to its experience by assuming physical state  $S_1$ . Distant particle B immediately assumes  $S_2$ . Here again, and for the same reasons, we don't have a violation of conservation laws. What about privacy and telepathy? We consider three cases. (1) B's physical nature responds physically to A's physical state  $S_1$ . Then B doesn't interact with A's experience and issues about privacy and telepathy don't arise. (2) B's physical nature responds to A's experience  $E_1$  by assuming physical state  $S_2$ . Here we can avoid problems of privacy by treating A and B as a single entity, the composite C.<sup>23</sup> Quantum

---

<sup>20</sup> See Lewtas (2013) for the kind of experiences A might have. That paper argues that possible complexity of experience depends on complexity of physical structure, notes that science reveals A as maximally simple, and accordingly judges A's experiences maximally simple — for example, an experience as of a borderless, homogeneous red which presents but doesn't represent.

<sup>21</sup> We should assume that A's physical nature behaves consistently — that the relations between  $E_1$  and  $S_1$  and between  $E_2$  and  $S_2$  are lawful. Otherwise we can't speak of A's *nature*.

<sup>22</sup> Seager (1999, p. 245) explores similar ideas.

<sup>23</sup> In this case, C has either a complex experience or a complex of experiences — in each case built of  $E_1$  and  $E_2$ . This shouldn't pose problems if, as in an earlier note, we tie possible complexity of experience to complexity of physical structure.

entanglement underwrites this move. This still leaves concerns about telepathy. Perhaps we can dodge these by regarding C as a single spread-out entity inside which everything relevant occurs. Or maybe we can claim that nonlocality is everyone's problem and therefore not especially ours. After all, physicists struggle with the faster-than-light transfers of influence/information that apparently occur in entangled systems. So we beg off until we better understand the physics. (3) B responds to events at A (involving either  $E_1$  or  $S_1$ ) by coming to experience  $E_2$ , to which B's physical nature responds by assuming physical state  $S_2$ . The same considerations apply. We needn't fret about privacy if, as above, we think of A and B as a composite. Then we don't have two separate consciousnesses — two separate subjects — but rather a single complex entity with a single complex consciousness.<sup>24</sup> And we can shelve concerns about telepathy, either by treating A and B as a single spread-out entity, or by begging off about non-locality.

We already discern the shape interactionist causation must take. Called here the *two-step trigger*, this typically involves two causal transactions and three components. It needs a conscious cause that doesn't act but just exists as is. It also needs a physical responder to respond lawfully to the conscious cause by assuming some physical state. Since this response mustn't violate any conservation laws, it shouldn't create, destroy, or transfer any conserved quantity.<sup>25</sup> The conscious cause and the physical responder together make up the first step and give the experience its first physical effect. Then, in typical cases but not necessarily all, some wider physical system responds, in standard physical ways, to the state assumed by the first physical responder. This second step does involve the transfer of conserved physical quantities. It enables the first step to serve as a cause for further and possibly large-scale physical effects.

Now for an example, followed by some remarks. Think of Schrödinger's cat. If an atom decays, a device in a box detects the decay and causes another device to break open a phial of poison. A

---

<sup>24</sup> For a rough analogy, think of a brain, part of which supports an experience of pain while another part supports an experience of red, with both feeding into an overall experience.

<sup>25</sup> The present paper ignores cases where the two-step trigger both creates and destroys a conserved quantity so as to maintain its overall amount — largely because they involve creation out of nothing and destruction into nothing. Lewtas (MS) argues against the possibility of such creations/destructions.

cat, imprisoned in the box, dies. If an atom doesn't decay, the first device detects nothing, the second device does nothing, and the cat lives. The box shows how a physical micro-event can trigger a physical macro-event. Now re-engineer the box. Put inside it our basic physical particle A. Rig up a first device that detects whether A occupies  $S_1$  or  $S_2$ , and then, if  $S_1$ , trips a second device that releases poison. Suppose, as before, that A comes to experience  $E_1$  and therefore to assume physical state  $S_1$ . The first device detects  $S_1$  and thus causes the second device to kill the cat. Note how A's experience  $E_1$  passively causes A's physical state  $S_1$  and, through  $S_1$ , the cat's death — but only because A's physical nature actively responds to  $E_1$ . The two devices, with energy sources and mechanisms independent of A, allow  $S_1$ , and thereby  $E_1$ , to cause the cat's death, a physical macro-event. Only because suitable physical mechanisms embed  $E_1$  can it bring about a macro-physical outcome.<sup>26</sup>

The present paper uses quantum physics solely to craft examples cleaving to actual science. We don't know whether consciousness has much to do with quantum processes. We have some reason to think not. Quantum events appear physically closed; quantum indeterminacy doesn't seem a promising foothold for determining conscious causes; and we have no evidence that consciousness skews quantum probabilities — although we have little direct evidence about quantum events in the brain.<sup>27</sup> The examples nevertheless highlight features needed for two-step triggers, features without which an experience can't passively cause a macro-physical event without violating conservation laws. So far science hasn't discovered non-quantum states possessing these features. The interactionist must therefore square consciousness with quantum physics or postulate some other physical entity/state/process that has the needed features. The paper follows the first course for expository purposes only.

We can now say something about the ability of passive causation to figure in extended causal chains. (1) We can have a causal chain leading from passive to active causation — where the conscious property passively causes a physical effect that in turn actively causes other physical effects. This follows from the actively causal/functional natures of all physical entities, including the passively-caused effect. (2) We can't have a chain leading from passive causation to passive

---

<sup>26</sup> The paper hasn't the space to address higher-level mental causation here.

<sup>27</sup> But see Penrose (1994) and McFadden (2000, chapter 13).

causation if, as argued earlier, physical entities have purely causal/functional natures. For the first passively-caused effect would lack a non-causal/non-functional qualitative nature that other physical entities could respond to. (3) We can have a chain leading from active causation to passive causation if physical entities can affect conscious properties. But this seems likely given the role of the physical base in the genesis of conscious emergents (on emergentism) and the role of physical structure in the construction of higher-level conscious properties (on particulate property dualism). (4) What about mental-to-mental causation? Chalmers' problem tells us that conscious properties can't have actively-causal natures. This means that mental-to-mental causation, if it exists, must take the form of passive causation. But passive causation requires the entity-suffering-the-effect to *actively* respond to the passive cause. The conscious entity-suffering-the-effect must therefore have an actively-causal nature — to respond to the passive conscious cause — and yet can't have an actively-causal nature — to comply with Chalmers' constraint. This apparently rules out mental-to-mental causal chains.

We now apply our model to standard emergentist dualism — where complex conscious properties emerge to characterize suitably organized physical composites. Suppose physical composite  $C_1$  — part of a brain, say — lawfully necessitates the emergence of complex experience  $E_1$  to characterize  $C_1$ . Interaction requires a physical responder  $P$  through which  $E_1$  has its first physical effect. We have to consider two cases: (A)  $P$  lies within and thus forms part of  $C_1$ ; and (B)  $P$  lies outside  $C_1$ . Consider (A) first. Here we run into trouble.  $C_1$  and  $E_1$  come into existence at the same time, since  $C_1$  necessitates  $E_1$ .<sup>28</sup>  $P$  responds to  $E_1$  at that very moment, since its constant nature

---

<sup>28</sup> The paper presupposes that emergence involves synchronic supervenience. Some deny this. O'Connor (2000) and O'Connor and Wong (2005) understand emergence as a non-supervenient, non-synchronic, dynamic relation — where supervenience fails, in part, because prior emergents contribute to the emergence base. Humphreys (1997) construes emergence as a non-supervenient, non-synchronic, dynamic process of fusion, wherein the base entities lose their individual identities and causal powers to fuse into a unified emergent with novel causal powers. (Humphreys argues that fusion overcomes concerns about causal exclusion since the emergent and its base entities never coexist.) Passive causation works on O'Connor's and Wong's account, which allows for qualitative conscious emergents. Indeed, the problem addressed in the text, which depends on synchronic supervenience, can't arise on their account. The status of passive causation on Humphreys' account is less clear. Humphreys requires emergent entities to have novel powers. If this means novel *active* powers, then Humphreys-style conscious emergents run afoul of Chalmers' problem. They also can't cause passively unless they

provides therefor. Now suppose  $E_1$  passively causes  $P$  to go from superposition of physical states  $S_1$  and  $S_2$  to determinate state  $S_1$ . But  $P$ -in-superposition, and not  $P$ -in- $S_1$ , forms part of  $C_1$ . It follows that  $C_1$  stops existing as soon as  $P$  responds to  $E_1$ :  $C_1$  changes into distinct physical composite  $C_2$ , exactly like  $C_1$  except with  $P$ -in- $S_1$  where  $C_1$  had  $P$ -in-superposition. Suppose also that  $E_1$  emerges only from  $C_1$ . Then, at one moment,  $C_1$  arises,  $E_1$  emerges,  $P$  responds,  $P$ -in-superposition becomes  $P$ -in- $S_1$ ,  $C_1$  ceases to exist,  $C_2$  arises, and  $E_1$  blinks out. Here we have the emergentist's problem of downward causation: the clash between the upward determination of experience by the physical and the downward causation of the physical by experience. Maybe the problem lacks bite: maybe  $E_1$  can make its mark despite a vanishingly brief existence. Still, a satisfactory theory would accommodate mental causation more securely. Not only might  $E_1$  need time to have effects, but also, introspection tells us that causally effective conscious states have temporal thickness.

Nature might solve this problem for us. Three possibilities come to mind. (A1)  $C_1$  continues to necessitate  $E_1$  notwithstanding small changes to it, including the shift in  $P$  from superposition of  $S_1$  and  $S_2$  to determinate state  $S_1$ . This makes some sense. After all, we would expect  $C_1$  to continue to necessitate  $E_1$  through slight temperature changes. (A2) A many-to-one supervenience relation obtains whereby distinct physical composites  $C_1$  and  $C_2$  both lawfully necessitate  $E_1$ .<sup>29</sup> Nothing prevents this: emergentism forbids a mental change without a physical change but not a physical change without a mental change. This yields the following:  $E_1$  emerges from and characterizes  $C_1$ ,  $E_1$  passively causes  $P$  (within  $C_1$ ) to assume physical state  $S_1$ , this turns  $C_1$  into  $C_2$ , but the shift to  $C_2$  has no effect on  $E_1$  because  $E_1$  also emerges from  $C_2$ . (A3)  $E_1$  does blink out when  $C_1$  turns into  $C_2$ , but  $P$  remains in determinate state  $S_1$  nonetheless. In this case  $E_1$  can make a causal difference notwithstanding a vanishingly short life. If any of these scenarios holds, then  $E_1$  can serve in a two-step trigger. We have no guarantee, of course, that any of them does. This shows that emergentist interaction occurs only if some additional condition obtains.

---

also have non-causal qualitative properties. The present author believes that the processes postulated by O'Connor, Wong, and Humphreys qualify neither as emergence nor as genuine metaphysical possibilities — claims impossible to defend here.

<sup>29</sup> (A2) might not differ from (A1).



Suppose that one of the three saving scenarios does hold. Suppose also that P lies in some larger physical structure  $C_3$  which intersects  $C_1/C_2$  at least at P. And allow that  $C_3$  ‘magnifies’  $S_1$  to give it macro-physical effects the same way the two devices inside the modified Schrödinger box ‘magnified’ micro-physical state  $S_1$  (from that example) so as to kill the cat.<sup>30</sup> Now we have emergent mental causation without wrinkles. The example doesn’t violate any conservation laws since  $E_1$  causes  $S_1$  without creating, destroying, or transferring any conserved physical quantity. And  $S_1$  causes other physical events courtesy of the physical quantities transferred within  $C_3$ . Nor does the example run afoul of privacy and telepathy.  $E_1$  emerges from  $C_1/C_2$  and has its passive effect on P within  $C_1/C_2$ . Everything happens within  $C_1/C_2$ .

Now consider case (B), where P lies outside  $C_1$  — and where, again,  $E_1$  emerges from  $C_1$  and passively causes P to assume physical state  $S_1$ . Allow, as before, that P lies within some physical composite  $C_3$  — perhaps disjoint from  $C_1$  — that ‘magnifies’  $S_1$  to give it wider macro-physical clout. Here we needn’t worry about  $C_1$ ’s becoming distinct composite  $C_2$  which might not lawfully necessitate  $E_1$  — because  $E_1$  here has its effect on something other than its supervenience base. But the example does run into trouble with privacy and telepathy. How does P, outside  $C_1$ , ‘reach out’ to  $E_1$  so as to respond to it? Three possibilities suggest themselves: (B1) P abuts  $C_1$ , (B2) some physical structure links  $C_1$  to P, and (B3) P and  $C_1$  are entangled (in something analogous to the quantum mechanical sense). (B1) doesn’t help. Contiguity either leaves P and  $C_1$  separate or else physically connects them. If it leaves them separate, it doesn’t touch the problem. And if it connects them, it puts us into case (B2). (B2) doesn’t solve the problem either. Any physical structure linking P and  $C_1$  in the required way would alter  $C_1$  so as to bring P within it. This takes us back to case (A) above. That leaves (B3), that P and  $C_1$  are somehow entangled. This does fix the problem, but at the cost of entangling two physical entities at least one of which is composite. But macro-physical entanglement is controversial science. We see that (B3) obliges the dualist either to anchor consciousness in quantum

---

<sup>30</sup> If P is itself a physical composite, then  $S_1$  might be a macro-physical state. Then  $C_3$  allows  $E_1$  other macro-physical effects, which, unlike the transition to  $S_1$ , do involve the transfer of conserved physical quantities.

processes, some of them speculative, or to postulate non-quantum entanglement.

This concludes our look at the mechanics of interactionism. Interestingly, the first kind of dualism — where basic particles have ‘atomic’ experiences — more easily accommodates interactionist causation. Emergentist dualism accommodates it only after satisfying further and, in some cases, controversial criteria.<sup>31</sup>

## 5. Functionalism

If dualist interaction does occur, then the functionalist understanding of mental causation doesn’t say enough. The functionalist story describes a two-tiered system with second-order functional properties and first-order physical realizers. Some functionalists identify conscious properties with second-order functional properties; others equate them with second-order functional properties as realized. Regardless, conscious properties inherit their causal powers from their physical realizers. The dualist can preserve the insights of functionalism by replacing this two-tiered model with a three-tiered model. The dualist insists that conscious properties don’t reduce to functional properties, realized or otherwise. He suggests instead that conscious properties — perhaps in tandem with the physical mechanisms that magnify or translate their effects — themselves realize functional properties. From this point of view consciousness carries out (at least some) mental functions. The dualist then proposes that physical entities *enable* these conscious properties. He has in mind, if an emergentist, the role of the subvenient physical in the genesis of conscious properties, or, if a particulate property dualist, the role of the basic physical particles which instantiate atomic experiences as well as the role of physical structures in combining these into complex experiences. The resulting three-tiered model provides that physical entities *enable* high-level but non-physical conscious properties which in turn (and perhaps in tandem with other physical entities) *realize* (at least some) mental functions. A simple example should serve. An instance of the physical property *being in such-and-such a neural state* might enable an instance of the non-physical property *being in*

---

<sup>31</sup> Emergent mental causation faces another problem. We saw that passive causation requires strong causation. But emergence requires weak causation. Hence emergent mental causation must conform to both strong and weak causation — a logical impossibility. See Lewtas (2017).

*pain*, which (perhaps in tandem with sundry physical entities) might realize an instance of the causal/functional property *detecting tissue damage*.

## 6. Consciousness Detectors?

Might interactionist mental causation allow for devices that detect consciousness, or even measure it, in apparent violation of the privacy of experience? In fact, not. All current scientific detectors suffer the effects of active physical causes. The relevant events all involve the transfer of conserved physical quantities. But a consciousness detector would need to respond actively to something essentially passive, without any transfer of a conserved physical quantity from passive cause to active detector. Suppose you built a machine of this sort. How would you know which passive state it had detected? You couldn't find out the usual way — by identifying the source of the conserved physical quantity affecting the detector. The problem becomes acute when we recall that experiences produce macro-physical effects (involving the transfer of conserved physical quantities) only through the activities of intermediate physical mechanisms. But then any given experience could cause any given macro-physical effect provided we placed between them the right intermediate mechanism. Hume complained that the connection between cause and effect was logically contingent albeit lawfully necessary. But, here, the connection between conscious cause and macro-physical effect is also lawfully contingent — until we specify an intermediate mechanism.<sup>32</sup> Suppose we overcome this hurdle by using trial-and-error to engineer a detector. That is, suppose we subject people to pain and then randomly fiddle with our machine — perhaps making massive changes to it — until the pain makes it light up (think of drug companies throwing a thousand chemicals at a pathogen in hopes of finding one that stops it). But what should we conclude if the machine doesn't light up while we tear the tentacles off an octopus? Would we have good cause to believe the octopus wasn't in pain?

We have another reason for thinking consciousness detectors unlikely. Each of the paper's examples featured a first physical responder inside the physical entity having the causally effective

---

<sup>32</sup> We assumed a lawfully necessary connection between basic physical particle A's experience  $E_1$  and its physical state  $S_1$ . Contingency returns when we embed A in a two-step mechanism giving  $E_1$  macro-physical efficacy.

experience — except for the example where the responder was entangled with the experiencing entity. Otherwise we would have had to postulate something like telepathy. This means that a consciousness detector would either have to become part of the entity tested — not just placed inside it, but made part of it — or else to become entangled with it. This might not be possible without changing, damaging, or even destroying the thing tested. Would you sign up for a scan?

### 7. Mental Causation or Not?

Should the dualist welcome this account? Not obviously. The last section pointed out that an experience can cause any macro-physical effect given the right intermediate mechanism. With suitable ‘wiring’, then, the experience of pain might lead one to seek and eat food. It seems to follow that evolution linked experiences and roles randomly. The dualist must find this troubling. Pain appears, by its nature, well suited to realize some roles and less well suited to realize others. Certainly most interactionists would want its hurtfulness to have something to do with its physical effects. But if it can’t, the dualist faces a watered-down epiphenomenalism: pain’s hurtfulness has macro-world effects, but the physical alone determines which. The dualist might now wonder why he struggled to give pain a causal role. Why shouldn’t he ditch interactionism, with its problems, and leave the physical in sole charge of making things happen?

Does the dualist have a good answer? Any reply must affirm that pain somehow does a better job realizing some roles than others. But how? Well, our account has said little about mental-to-mental causation — where experience  $E_1$  directly causes experience  $E_2$  — except to dismiss it as impossible. But this still leaves two other ways mental-to-mental causation might occur: (1) through higher-level causation, about which the paper has said nothing; or (2) through causal chains, wherein a mental entity passively causes a physical entity, which, directly or via some intermediate physical-to-physical causal chain, *enables* some experience. Suppose one of these furnishes a workable model of mental-to-mental causation.<sup>33</sup> And suppose — a big if — that the successful model allows pain only ‘common-sense’ mental effects: a desire to end the pain, a desire to move away from the

---

<sup>33</sup> The paper hasn’t the space to explore the relevant issues here.

source of tissue damage, etc., and not a desire to seek and eat food.<sup>34</sup> The dualist might then say that all these experiences, through their causal effects on one another, hold together as a structured whole. He might have in mind something like a functionalist machine description — where experiences occupy nodes connected by directed (one-way or two-way) causal links. The resulting network has a structure, a shape, determined by the number of nodes and the pattern of linkages between. The dualist might then maintain that this shape permits the organism effective traffic with its environment only when the individual experiences have common-sense physical effects — when pain causes the organism to move away from sources of tissue damage and hunger causes it to seek and eat food. After all, the network of experiences must hook up with sensory or other inputs at the receiving dock and with behaviours or other outputs at the shipping bay. And its shape will affect its capacity to do this. Suppose that evolution accords pain the role of pushing an organism to seek and eat food rather than the role of pulling it away from sources of tissue damage. And assume that pain's hurtfulness bears commonsensically on its mental effects. The structure of experiences generated by the pain, through its common-sense mental effects, has a certain shape. And the structure of experiences needed to cause the behaviours appropriate to seeking and eating food has its shape. Maybe the network of experiences generated by pain has only two output causal links — the number apt for causing organisms to move away from sources of tissue damage. And maybe it takes three output causal links to cause the full suite of behaviours involved in seeking and eating food. This yields a mismatch between the shape the network has and the shape it needs in order to occasion the appropriate behaviours — a mismatch reflected in more or less unfit behaviour. The greater the mismatch between the experiences' mental effects and physical effects, the more we should expect the network's shape to differ from the shape it needs in order to cause the relevant behaviours. At some point the network won't be able to cause the behaviours. Under these conditions, evolution will tend to settle on the 'right' links between experiences and their physical effects.

The dualist must do much to flesh out this picture. First, he owes us a fully elaborated account of mental-to-mental causation that grants

---

<sup>34</sup> The dualist thus claims that experiences relate necessarily to their mental effects even though contingently to their macro-world physical effects.

mental causes only common-sense mental effects. Second, he must demonstrate that some structures of experiences couldn't effectively cause certain structures of behaviours (given the chosen theory of mental-to-mental causation). The first task might not prove as daunting as it appears. For the original problem — that a mental cause can have any physical macro-effect given the right intermediate mechanism — haunts any account of mental causation, even those physicalist accounts crafted to preserve the causal relevance of mental entities. Most of these, after all, have the mental cause (e.g. the decision to jump) causing a physical effect (the springing of legs) via a chain of intermediate physical causes (neural and muscular events in brain, trunk, and legs). This suggests the problem might have a solution now unknown. Failing that, a problem confronting everyone shouldn't reproach interactionism in particular.

### 8. Conclusion

The present paper hasn't proved — hasn't even tried to prove — that interactionist causation occurs. It has instead asked what shape it would have to take if it did occur, and what implications this might have for the wider concrete world. Throughout it has sought an account consistent with both non-physicalist property dualism and our best science. The paper aimed for something else, too. Many philosophers deride dualism as essentially negative and therefore sterile. Dualism denies that consciousness is physical, they claim, without supplying an alternate research programme, much less a promising one. Perhaps some of the paper's arguments and models hint at responsible and fruitful ways of doing dualist philosophy.<sup>35</sup>

### References

- Armstrong, D.M. (1983) *What is a Law of Nature?*, Cambridge: Cambridge University Press.
- Averill, E. & Keating, B. (1981) Does interactionism violate a law of classical physics?, *Mind*, **90**, pp. 102–107.
- Bird, A. (2007) *Nature's Metaphysics: Laws and Properties*, Oxford: Clarendon Press.
- Chalmers, D. (1996) *The Conscious Mind*, New York: Oxford University Press.
- Crane, T. (1995) The mental causation debate, *Proceedings of the Aristotelian Society*, suppl. vol. **69**, pp. 211–236.

---

<sup>35</sup> The author thanks two reviewers — for this journal and another — for extremely helpful comments.

- Davidson, D. (1995) Thinking causes, in Heil, J. & Mele, A. (eds.), *Mental Causation*, pp. 1–17, Oxford: Clarendon Press.
- Dretske, F. (1977) Laws of nature, *Philosophy of Science*, **44**, pp. 248–268.
- Eccles, J. (1987) Brain and mind: Two or one?, in Blakemore, C. & Greenfield, S. (eds.) *Mindwaves*, Oxford: Blackwell.
- Ellis, B. (2001) *Scientific Essentialism*, Cambridge: Cambridge University Press.
- Ellis, B. (2002) *The Philosophy of Nature*, London: Acumen.
- Gibb, S. (2013a) Introduction, in Gibb, S., Lowe, E. & Ingthorsson, R. (eds.) *Mental Causation and Ontology*, pp. 1–17, Oxford: Oxford University Press.
- Gibb, S. (2013b) Mental causation and double prevention, in Gibb, S., Lowe, E. & Ingthorsson, R. (eds.) *Mental Causation and Ontology*, pp. 193–214, Oxford: Oxford University Press.
- Heil, J. (2003) *From an Ontological Point of View*, Oxford: Clarendon Press.
- Humphreys, P. (1997) How properties emerge, *Philosophy of Science*, **64**, pp. 1–17.
- Jackson, F. (1982) Epiphenomenal qualia, *Philosophical Quarterly*, **32**, pp. 127–136.
- Kim, J. (1995) Can supervenience and ‘non-strict laws’ save anomalous monism?, in Heil, J. & Mele, A. (eds.) *Mental Causation*, pp. 19–26, Oxford: Clarendon Press.
- Kim, J. (2000) *Mind in a Physical World*, Cambridge, MA: MIT Press.
- Kim, J. (2005) *Physicalism, or Something Near Enough*, Princeton, NJ: Princeton University Press.
- Kim, J. (2010) *Philosophy of Mind*, 3rd ed., Boulder, CO: Westview Press.
- Larmer, R. (1987) Mind–body interactionism and the conservation of energy, *International Philosophical Quarterly*, **26**, pp. 277–285.
- Lewis, D. (1973) Causation, *Journal of Philosophy*, **70**, pp. 556–567.
- Lewis, D. (1980/1990) What experience teaches, in Lycan, W. (ed.) *Mind and Cognition*, Oxford: Basil Blackwell.
- Lewtas, P. (2013) What it is like to be a quark, *Journal of Consciousness Studies*, **20** (9–10), pp. 39–64.
- Lewtas, P. (2017) The impossibility of emergent conscious causal powers, *Australasian Journal of Philosophy*, **95** (3), pp. 475–487.
- Lewtas, P. (MS) Conservation of almost everything.
- Lowe, E. (2008) *Personal Agency: The Metaphysics of Mind and Action*, Oxford: Oxford University Press.
- Lowe, E. (2013) Substance, causation, powers, and human agency, in Gibb, S., Lowe, E. & Ingthorsson, R. (eds.) *Mental Causation and Ontology*, pp. 153–172, Oxford: Oxford University Press.
- Mackie, J. (1980) *The Cement of the Universe*, Oxford: Oxford University Press.
- Martin, C.B. (2008) *The Mind in Nature*, New York: Oxford University Press.
- McFadden, J. (2000) *Quantum Evolution*, New York: W.W. Norton & Company.
- Mills, E. (1996) Interactionism and overdetermination, *American Philosophical Quarterly*, **33**, pp. 105–115.
- Molnar, G. (2003) *Powers: A Study in Metaphysics*, New York: Oxford University Press.
- Mumford, S. (1998) *Dispositions*, Oxford: Oxford University Press.
- Mumford, S. (2004) *Laws in Nature*, Abingdon: Routledge.
- Mumford, S. (2009) Passing powers around, *The Monist*, **92** (1), pp. 94–111.

- Mumford, S. & Anjum, R. (2010) *Getting Causes from Powers*, Oxford: Oxford University Press.
- O'Connor, T. (2000) Causality, mind and free will, *Philosophical Perspectives*, **14**, pp. 105–117.
- O'Connor, T. & Wong, H. (2005) The metaphysics of emergence, *Nous*, **39**, pp. 658–678.
- Penrose, R. (1994) *Shadows of the Mind*, Oxford: Oxford University Press.
- Prior, E., Pargetter, R. & Jackson, F. (1982) Three theses about dispositions, *American Philosophical Quarterly*, **19** (3), pp. 251–257.
- Psillos, S. (2002) *Causation and Explanation*, Montreal: McGill and Queen's University Press.
- Seager, W. (1999) *Theories of Consciousness*, New York: Routledge.
- Shoemaker, S. (1980) Causality and properties, in Van Inwagen, P. (ed.) *Time and Cause*, Dordrecht: Reidel.
- Stapp, H. (1993) *Mind, Matter, and Quantum Mechanics*, Berlin: Springer-Verlag.
- Strawson, G. (2008) David Hume: Objects and power, in *Real Materialism and Other Essays*, Oxford: Oxford University Press.
- Strawson, G. (2016) Mind and being: The primacy of panpsychism, in Brüntrup, G. & Jaskolla, L. (eds.) *Panpsychism: Contemporary Perspectives*, Oxford: Oxford University Press.
- Tooley, M. (1987) *Causation*, Oxford: Clarendon Press.
- Williams, N. (2011) Dispositions and the argument from science, *Australasian Journal of Philosophy*, **89**, pp. 79–90.

Paper received September 2017; revised March 2018.