The Possibility of Emergent Conscious Causal Powers

Lok-Chi Chan, Andrew James Latham

[This is a preprint of an article whose final and definitive form is published in the *Australasian Journal of Philosophy*. The *Australasian Journal of Philosophy* is available online at: <http://www.tandf.co.uk/journals/>. Please cite the published version.]

**Abstract:** Lewtas [2017] recently articulated an argument claiming that emergent conscious causal powers are impossible. In developing his argument, Lewtas makes several assumptions about emergence, phenomenal consciousness, categorical properties, and causation. We argue that there are plausible alternatives to these assumptions. Thus, the proponent of emergent conscious causal powers can escape Lewtas’s challenge.

**Keywords:** emergentism; laws of nature; mental causation; qualia; categorical properties

**1. Introduction**

Lewtas [2017] recently articulated an argument for the impossibility of emergent conscious causal powers. For him, there are two possible accounts of such powers. Either emergent consciousness brings about its effects ‘actively’ or it brings it about ‘passively’. Lewtas presents his argument as a dilemma with each account being problematic for different reasons.

There are many accounts of emergence in the literature. Weak versions of emergence are compatible with the standard formulations of physical realization,[[1]](#footnote-2) according to which the realized property is not distinct from its realizer properties but is the result of their joint contributions. Lewtas, though, is concerned with a strong version of emergence, according to which the emergent phenomena and physical bases are wholly distinct. He describes the emergent phenomena as ‘a sui generis mental force, something akin to telekinesis’ [Lewtas 2017: 478]. We agree with Lewtas that there are no strongly emergent conscious causal powers, but disagree they are impossible. The emergentist can escape Lewtas’s challenge by rejecting his metaphysical assumptions.[[2]](#footnote-3)

Following Lewtas’s treatment of active and passive causation, our discussion of the two kinds of causation is independent from one another. Our discussion of active causation focusses on Lewtas’s characterizations of science, the hard problem of consciousness, and the metaphysics of properties. Meanwhile, our discussion of passive causation focusses on Lewtas’s analysis of the metaphysics of causation.

**2.** **Active causation**

According to the active account, phenomenal consciousness brings about an effect through *its own* causal efficacy. In developing his counterargument, Lewtas assumes that the physical should be defined by the posits of physical science and that science can only ever ‘gather data about manifested dispositions’ and not intrinsic, categorical properties [478]. Hence, everything physical is dispositional in nature [478].

Lewtas’s argument is as follows. The emergentist must make a further assumption about the nature of phenomenal consciousness: ‘the emergentist can treat consciousness as non-physical only if he grants it a non-causal . . . nature’ [478]. If this nature is not granted, then it follows that physical properties can play the role of consciousness. So, physicalism is a better theory because it is simpler and more uniﬁed [479]. If phenomenal consciousness is a categorical property and possesses causal powers, then it has both a categorical and a causal component. But, since ‘nothing essentially non-causal . . . can equal anything essentially causal’ [479] the emergentist must ‘recognize them as two separate entities, with disparate natures, somehow wrapped together’ [479].

It is unclear why the emergentist should accept this argument. Emergentism is the view that there are some high-level novel properties which are irreducible to, and unpredictable in terms of, the fundamental properties. There might be high-level causal properties of this kind, and even if physicalist explanations of these properties are possible, this does not automatically mean they are the best ones. Simplicity and unification are not the only considerations when choosing theories, especially when different camps have different views on what considerations—theoretical virtues, explananda, and the like—there are, and which are most important.

Lewtas though highlights an important point. A common response to emergentism is that failures of reductive accounts might be due to incomplete scientific knowledge, and emergentists cannot rule this out. However, emergentist accounts of consciousness place consciousness into an ontological category *wholly distinct* from that of physical properties, making it exempt from physical explanations. But even if the emergentist adopts this strategy, it is unclear why she can only choose the non-casual category. Emergentists and other non-physicalists have appealed to numerous other ontological categories in the past. For instance: (1) being intentional (for example, Kripke [1982] and Bealer [1997]), (2) being first-person or subjective (for example, Nagel [1974]), (3) being self-aware (for example, Bealer [1997]), (4) being the basis of agent causation (for example, O’Connor [2000]), and (5) being irreducibly macroscopic (for example, Van Cleve [1990] and James [1890]). Importantly, none of these ontological categories is obviously in conflict with being causal, nor excludes the possibility of possessing active powers.

One related worry is that the non-casual category might not be wholly distinctfrom the physical. Lewtas here follows metaphysicians who defend ‘categorical humility’—the view that physical science can never know of the non-casual, intrinsic categorical bases of things ([Russell 1912; Langton 1998]; see also Lewtas [2017: 478]). We take no stand here on categorical humility, except to say that it is not obviously true when considered alongside scientific theories.

There is the possibility that physical science can posit non-casual properties. For instance, physicists posit what they call ‘intrinsic properties’ like mass: mass is not a force; it is a property that *correlates* with inertia and gravitational forces by playing the roles of inertial mass, active gravitational mass, and passive gravitational mass. While humility theorists believe that mass is a causal-dispositional property, others have suggested that it might be an intrinsic, categorical property (see, for example, Smart [1963], Ney [2015], and Hiddleston [2019]). For these philosophers, we *discover, define, and measure* an object’s mass (a quantitative nature intrinsic to that object) via observing the object’s dispositions, but it is just *revealed* by, and is not *identical* with, the relevant dispositions [Smart 1963: 74].

Of course, humility theorists believe that scientific inquiry is restricted to the casual, though there is no agreement why this is the case. We will set aside arguments for humility and instead demonstrate our point by diagnosing the problem with Lewtas’s argument. Lewtas first appeals to a widely accepted view that ‘all data must causally affect scientists’ senses, either directly, or indirectly through prior effects on instruments’ [Lewtas 2017: 478]. He then argues that since categorical properties cannot casually affect us, they cannot be known to us (other formulations of this argument are in Russell [1912] and Langton [1998]).

This argument is problematic because it reflects a narrow, a priori conception of science according to which the explanatory relevance of something can only be grounded in its causal connections to the explanandum. This view has been rejected by some scientists (for example, Carroll [2015: 126]), for not all scientific explanations are causal explanations. There are, for instance, synchronic explanations in which we explain a phenomenon by positing the entities that ground it. While it is natural to think that the posited entity must bear some ontological relation to the explanandum, the relation need not be a causal one. Instead, it could be one of, amongst others, realization, determination, necessitation, supervenience, composition, or structural support, none of which must be causal. Physicist Sean Carroll [2015: 126] uses the multiverse theory—which posits many universes that are causally disconnected from us—as an example to demonstrate this point. This means that we must loosen the a priori causal restriction on science which Lewtas assumes. However, when we do this, we can no longer rule out categorical properties from being explanatory. We might even be able to grasp them, just like other entities uncovered by science.

What does this have to do with Lewtas’s discussion of emergence? There are two implications. The first concerns the nature of consciousness and why the hard problem of consciousness seems persistent. As previously noted, Lewtas argues that, for the emergentist, consciousness should be understood as non-physical because it is a kind of categorical property, whereas physical properties should be understood as causal properties. But we now see that some physical properties, like mass, could plausibly be categorical properties. So Lewtas’s argument, which appeals to the distinction between categorical and causal properties, is contentious.

Of course, no one thinks the hard problem of consciousness can be solved by appealing to something like categorical mass properties. But this intuition merely shows that the reason why the hard problem is hard is, perhaps, not due to consciousness being a non-causal categorical property, but rather due to something else. After all, we have intuitions regarding the uniqueness of consciousness, but these intuitions do not reveal precisely *why* consciousness is unique. This supports our earlier speculation that some other unique nature of consciousness (for example, being irreducibly macroscopic) might be crucial to why the hard problem is hard. While these options remain open, Lewtas’s argument does not generalize to them.

Secondly, Lewtas argues that categorical and causal properties are non-identical, and so are ‘separate entities, with disparate natures’ [Lewtas 2017: 479]. This argument is also contentious. As Oderberg remarks, two properties can be inseparable due to their natures, but this inseparability does not imply identity. For instance, the radius of a circle is inseparably related to its circumference while not being identical to it [Oderberg 2009: 678]. Of course, the relation between categorical and dispositional properties is more sophisticated than this. However, Oderberg’s point suffices to show that there can be necessary connections between non-identical properties. Consider again Smart’s view on the relation between mass and gravitation, a possible example of Oderberg’s point in the metaphysics of science. Mass and gravitation are inseparably related due to their natures or some law of nature, but nevertheless are non-identical. Similarly, for the emergentist, pain and pain-related psychology and behaviour, for example, might be inseparably related due to their natures or some law of nature.

To conclude, Lewtas’s argument against active causation relies on particular characterizations of science, the hard problem of consciousness, and the metaphysics of properties. We have argued that categorical humility as part of science is contentious, since there plausibly are categorical properties posited by non-causal explanations. Two implications are as follows. First, framing the hard problem in terms of phenomenal consciousness as non-causal categorical properties is questionable. Second, even if phenomenal consciousness is a kind of categorical property, given that categorical and causal physical properties might be inseparably connected, phenomenal consciousness too might be connected with its powers in this way.

These two solutions characterize phenomenal consciousness differently, but either way will defuse Lewtas’s argument. Both solutions also share the distinct benefit for emergentists of making phenomenal consciousness scientifically accessible. Phenomenal consciousness is either not a categorical property, or a categorical property available to non-causal scientific explanations. Hence, the strong emergentist project could be naturalistic and empirically tractable.[[3]](#footnote-4)

**3.** **Passive causation**

Let us move to Lewtas’s argument against the passive account of emergent conscious causal powers. On this account, phenomenal consciousness, by its own nature, has no causal powers; instead, physical entities actively respond to it.

In his argument, Lewtas assumes that emergence is a kind of causation, because it involves ‘a genuine change through time—the coming into being of the emergent out of nowhere and from nothing’ [Lewtas 2017: 482]. We find this assumption unusual, for it seems to conflate together two different kinds of causation: (1) *diachronic causation* which occurs between states of affairs at different times, and (2) *metaphysical causation* which runs along the axis of fundamentality, from most fundamental to least fundamental, at a time. Some philosophers will reject there is such a thing as metaphysical causation because they do not think that entities across levels are wholly distinct. But recall that the kind of emergentism discussed by Lewtas takes the emergent and its physical bases to be wholly distinct. So Lewtas might be right: some metaphysical relation, such as causation, might be required to connect these wholly distinct entities. Of course, it is up for grabs what that relation might be. Even if one disagrees with Lewtas that the right relation is causation, one might instead consider the recent literature on metaphysical grounding, which is a generative inter-level metaphysical relation. Some authors in that literature take both diachronic causation and inter-level metaphysical relations to be generative relations of the same kind [Schaffer 2016; Bennett 2017; Wilson 2018]. All that matters for our purposes here is to note that one may consider Lewtas’s argument as being concerned with generative relations, as opposed to anything specific about causation.

Here then is Lewtas’s argument. He begins by characterizing three theories of causation:

**The Humean account:** Causation is merely the arrangements of entities.

**The governing law account:** Some laws of nature which are irreducible, non-concrete entities govern the causal behaviors of concrete entities.

**The power account:** Entities themselves have powers essentially and thereby necessitate how they themselves causally behave.

For Lewtas, emergence is incompatible with the power account because the latter implies reductionism. Specifically, ‘emergentists must deny more-than-merely-correlational connections between base and emergent’ [482], or else the emergent would be ‘nothing over and above’ its bases [477]. But under the power account, the cause itself has the power to bring about the effect, and this entails the aforementioned more-than-merely-correlational relation.

However, the existence of passive causation also entails the power account. For only when some entities have powers in their nature, can there be active causal powers, and not mere correlations. Otherwise, it would ‘rule out even the possibility of distinguishing passive from active causes, and therewith even the possibility of recognizing passive causation as a distinct kind’ [483]. But both the Humean account and the governing law account involve no active powers: the former account involves only mere regularities, and under the latter account entities are connected by laws of nature that are extrinsic to them. In sum, because passive causation both requires and is incompatible with the power account, it is conceptually impossible.

We think this argument is unsound. It assumes that metaphysical causationand diachronic causation can both be accounted for by the *same* *theory* of causation. But if there are really two different kinds of causation, why assume that they share the same nature? While different accounts of causation are incompatible when they are about the same causal process, that incompatibility is no longer obvious if they are about *different kinds of* causal processes.[[4]](#footnote-5) Perhaps diachronic causation fits the description of the power account, whereas metaphysical causation fits the description of the Humean account or the governing law account. If this is right, then the emergentist can offer Lewtas everything he asks for: both ‘merely-correlational connections between base and emergent’ and active and passive causal powers. Specifically, the emergent is merely correlated with its physical bases, but the physical nevertheless possesses power and actively responds to the emergent.

One benefit for the emergentist for being a pluralist about causation is that it makes sense of the difference we observe between diachronic and synchronic cases. For instance, diachronic cases typically involve matters like energy transference [Braddon-Mitchell 1993; Dowe 2000], whereas no such matters are involved in synchronic cases. Of course, many metaphysicians think that only one account of causation holds true. This is understandable because the relevant accounts are typically conceived of as rivals, and many emphasize the importance of simplicity and unity when doing metaphysics. But if there are, contra those metaphysicians’ assumption, two kinds of causation instead of one, then it seems natural that we might need a different account for each of them. Further, this view is not lacking in unity, as there is only one account for each kind of causation. Finally, we so far have followed Lewtas’s assumption that metaphysical causation, like diachronic causation, really counts as causation. However, the point remains the same if we instead consider both relations as being generative relations: there might be different generative relations with different metaphysical natures.

Throughout our discussion, we have suggested several alternatives to Lewtas’s assumptions that do not rule out emergent conscious powers. The emergentist may appeal to them to hold on to emergentism. One might ask why we reject emergent conscious powers if we find them possible. The reason is that there is, from our perspective, no empirical ground for supposing they actually exist [Chan and Latham 2019]. But we doubt that they can be rendered impossible from the armchair. We hope too that our discussion of emergence can supplement some accounts of emergence outside of consciousness studies, such as those in quantum mechanics, though this is work for another discussion.

Acknowledgements

We thank Eran Asoulin, Duen-min Deng, Alexandre Erler, Graham Oppy, Luke Roelofs, and Wai-hung Wong for their feedback.

Funding

Chan acknowledges the support from the Ministry of Science and Technology of Taiwan (grant number: 108-2410-H-002-241-MY3). Latham acknowledges the support from the Ngāi Tai Ki Tāmaki Tribal Trust.

REFERENCES

Aizawa, Kenneth and Gillett, Carl 2009. The (Multiple) Realization of Psychological and Other Properties in the Sciences, *Mind and Language* 24/2: 181–208.

Balaguer, Mark 2004. A Coherent, Naturalistic, and Plausible Formulation of Libertarian Free Will, *Noûs*, 38/3: 379–406.

Bealer, George 1997. Self-Consciousness, *The* *Philosophical Review*, 106/1: 69–117.

Bennett, Karen 2017. *Making Things Up*, Oxford: Oxford University Press.

Braddon-Mitchell, David 1993. The Microstructural Causation Hypothesis, *Erkenntnis*, 39/2: 257–283.

Carroll, Sean 2015. Falsifiability, in *This Idea Must Die*, ed. John Brockman, New York: HarperCollins Publishers: 124–127.

Chan, Lok-Chi and Latham, Andrew James 2019. Four Meta-methods for the Study of Qualia, *Erkenntnis*, 84/1: 145–167.

Dowe, Phil 2000. *Physical Causation*, Cambridge: Cambridge University Press.

Hiddleston, Eric 2019. Dispositional and Categorical Properties, and Russellian Monism, *Philosophical Studies*, 176/1: 65–92.

James, William 1890 (1918). *The Principles of Psychology*, Vol. 1, New York: Henry Holt and Company.

Kripke, Saul 1982. *Wittgenstein on Rules and Private Language: An Elementary Exposition*, Cambridge, MA: Harvard University Press.

Langton, Rae 1998. *Kantian Humility: Our Ignorance of Things in Themselves*, Oxford: Oxford University Press.

Lewtas, Pat 2017. The Impossibility of Emergent Conscious Causal Powers, *Australasian Journal of Philosophy*, 95/3: 475–487.

Lowe, E. J. 2008. *Personal Agency: The Metaphysics of Mind and Action*, Oxford: Oxford University Press.

Nagel, Thomas 1974. What Is It Like to Be a Bat?, *The Philosophical Review*, 83/4: 435–450.

Ney, Alyssa 2015. A Physicalist Critique of Russellian Monism, in *Consciousness in the Physical World: Perspectives on Russellian Monism*, eds. Torin Alter and Yujin Nagasawa, New York: Oxford University Press: 346–369.

O’Connor, Timothy 2000a. *Persons and Causes: The Metaphysics of Free Will*, New York: Oxford University Press.

Oderberg, David 2009. The Non-Identity of the Categorical and the Dispositional, *Analysis*, 69/4: 677–684.

Piccinini, Gualtiero and Maley, Corey 2014. The Metaphysics of Mind and the Multiple Sources of Multiple Realizability, in *New Waves in Philosophy of Mind*, eds. Mark Sprevak and Jesper Kallestrup,Bashingstoke: Palgrave Macmillan: 125–152.

Russell, Bertrand 1912 (1978). *The Problems of Philosophy*, Oxford: Oxford University Press.

Schaffer, Jonathan 2016. Grounding in the Image of Causation, *Philosophical Studies*, 173/1: 49–100.

Smart, J. J. C. 1963. *Philosophy and Scientific Realism*, London: Routledge.

Van Cleve, James 1990. Mind–Dust or Magic? Panpsychism versus Emergence, *Philosophical Perspectives*, 4: 215–226.

Wilson, Alastair 2018. Metaphysical Causation, *Noûs*, 52/4: 723–751.

1. For these formulations of physical realization, see, e.g., Aizawa and Gillett [2009]; Piccinini and Maley [2014]. [↑](#footnote-ref-2)
2. We follow the target article in setting aside Kim’s exclusion argument against emergentism [Lewtas 2017: 476]. [↑](#footnote-ref-3)
3. For notable examples of naturalistic emergentist projects, see O’Connor [2000] and Balaguer [2004]. [↑](#footnote-ref-4)
4. Lowe’s [2008] work on emergent mental causation involves a similar claim, albeit differently motivated. [↑](#footnote-ref-5)