



Consciousness, Causation, and Confusion

Darryl Mathieson^{1,2} 

Accepted: 11 November 2024

© The Author(s), under exclusive licence to Springer Nature B.V. 2024

Abstract

In recent decades, various sources of scientific evidence have been used to disprove the idea that we have free will. Conceptual confusion has tended to lurk behind such claims, however, for it often turns out that what researchers mean by “free will” is something extremely implausible. A similar problem persists in the literature on the causal role of consciousness. Various sources of psychological evidence have been used to show that consciousness is not among the causes of our behaviour. However, consciousness is often assumed to be both non-physical and non-functional, and causation is often poorly understood or not defined at all. The aim of this paper is to clear up these issues by urging that (1) future research on this topic remains neutral on the metaphysics of consciousness, (2) avoids assuming that phenomenal consciousness is functionless, (3) is more precise in specifying what is doing the causal work when discussing conscious mental states, and (4) adopts the practice of being clearer on what it takes for something to count as a cause. Much unwarranted scepticism about whether consciousness is causally efficacious has resulted from this fragmented thinking, and much is to be gained by clearing it up.

1 Introduction

It is well known that most of what goes on inside the human mind is unconscious.¹ This can even include very sophisticated activities and actions like reasoning, making inferences, pursuing goals, doing arithmetic, making decisions, thinking,

¹ By “conscious” and “unconscious”, I am referring to mental states in the following two distinct senses. The first is *phenomenal consciousness*. Briefly, a mental state is phenomenally conscious if there is something it feels like to be in that state (Nagel 1974). The second is *access consciousness*. This refers to states that are poised for use in reasoning and rationally controlling action and speech, for which reportability is a common marker (see Block 1995). I will move between the two throughout, given that the researchers whose work I engage with often differ in the senses they operate with. Nevertheless, I will try to indicate the relevant senses where appropriate.

✉ Darryl Mathieson
Darryl.Mathieson@outlook.com

¹ Australian National University, Canberra, ACT, Australia

² University of Cambridge, Cambridge, United Kingdom

self-control, and learning.² Building off this impressive literature, it has even been argued that unconscious processes can perform all the fundamental high-level functions that were traditionally thought to require consciousness.³ Recently, the journal *Frontiers in Psychology* dedicated a whole research topic to “The Epiphenomenalist Challenge” posed by such psychological research, summarized by one contributor as follows: “Is epiphenomenalism virtually entailed by the current empirical knowledge about how the mind/brain causes human behaviour?” (Lavazza 2019, p. 1).⁴ According to some psychologists, the answer is an emphatic *yes* (Halligan and Oakley 2021; Oakley and Halligan 2017; Pockett 2004).

But has science really shown that consciousness is a causally inert epiphenomenon? This would be a very substantive—and perhaps devastating—finding, given the central role that consciousness seems to occupy in human thought and action. Much has already been written (and found wanting) about the specific psychological evidence marshalled in support of epiphenomenalism (see Balaguer 2019; Baumeister et al. 2018; Lavazza 2019; Mathieson 2024; Mele 2009, 2018; Nahmias 2002, 2011; Stockdale 2022). However, much less attention has been paid to exploring more carefully what researchers seem to have in mind by consciousness and causation. If it turns out that the concepts of consciousness and causation are being defined and deployed in questionable ways—as I believe they are—then this is a further problem worth identifying and clearing up in responding to the epiphenomenalist challenge.

Such scepticism about the questionable assumptions underlying the concepts of consciousness and causation is not without good precedent. There has been a flurry of claims both in popular media and scientific journals in recent decades about science having undermined free will, only for it to turn out that what is meant by “free will” is something extraordinarily implausible (like an uncaused cause). I suspect a similar problem is going on in some of the literature regarding the causal efficacy of consciousness. Diagnosing the various issues and offering more plausible ways to understand consciousness and causation helps us to get a clearer idea of what is presently muddying the water in the epiphenomenalist challenge, and dovetails nicely with the vast existing work that has mainly been preoccupied with debating the methodology and interpretations of various experimental results.

This paper will proceed as follows. In § 2, I outline a problem with some of the scientific debunking of free will and argue that a similar issue persists in the consciousness literature on epiphenomenalism. In § 3, I address two issues generating confusion about how consciousness could be among the causes of behaviour: the assumption that consciousness is non-physical, and the denial that consciousness is a functional notion. In § 4, I address two issues generating confusion about causation in the consciousness literature: ambiguity in what is meant by expressions like “conscious thoughts”, and a lack of clarity about what exactly it is for something to count as a cause. I offer some brief concluding remarks in § 5.

² See Bargh (2017), Bargh and Williams (2006), Bargh and Morsella (2008), Dijksterhuis and Nordgren (2006), Libet (1985), Reber and Allen (2022), Soon et al. (2008), Wegner (2002), Wilson (2002).

³ For some recent examples, see Hassin (2013), Goldstein and Hassin (2017), Sklar and Hassin (2014), Sklar et al. (2021). For criticisms, see Hesselmann and Moors (2015) and Mathieson (2024).

⁴ Epiphenomenalism in the present context is the thesis that consciousness is caused by (or bears some other important relation to) physical brain processes but is itself entirely causally inert.

2 Science, Free Will, and the Causal Efficacy of Consciousness

Our actions have causal histories. Many have worried that what we know about these causal histories tells us that our actions are not free.⁵ Sometimes the worry is that key parts of these histories concern matters we have, at the time of acting, no control over; sometimes the worry is that determinism is true for our actions, whatever may be the case at the micro-level; sometimes the worry is prompted by one or another discovery about the role of certain physical processes in our brains in causing what we do; and so on and so forth. Wherever you stand in this debate, one thing seems clear and is widely acknowledged: it is important to spend time getting clear about the concept of a free action. It is easy to establish that something does not exist by defining it unrealistically, and a non-trivial proportion of the scientific literature on free will amounts to exactly that.

For example, Montague (2008) writes: “Free will is the idea that we make choices and have thoughts independent of anything remotely resembling a physical process. Free will is the close cousin to the idea of the soul — the concept that ‘you’, your thoughts and feelings, derive from an entity that is separate and distinct from the physical mechanisms that make up your body” (p. 1). In a recent (and much discussed) book, Sapolsky (2023) writes: “Show me a neuron (or brain) whose generation of a behaviour is independent of the sum of its biological past, and for the purposes of this book, you’ve demonstrated free will” (p. 15). Consider too the following take by Haggard (2005):

Several researchers have sought to relate the brain’s preparation of action to the philosophical concept of ‘free will’. Descartes proposed that the mind selects between alternative actions, and then causes the body, via the brain, to perform the selected action. This concept of action is deeply embedded in many modern societies, and is a key part of our folk psychology. However, it is incompatible with modern neuroscience, because it is strongly dualist, and implies mind– body causation (pp. 290–291).

These are absurdly high standards for what it means to have free will, and it is certainly not a reasonable starting point to *define* free will in these ways. *Of course* we do not act, make choices, or think independently from physical processes. *Of course* neurons, brains, and behaviour do not operate or occur independently of previous events and processes in the causal chain of our lives. If free will means any of that, then clearly, we do not have it. But it does not follow that we lack free will in some more realistic sense. Being sufficiently responsive to reasons for action would be one obvious candidate, and this is quite compatible with everything having antecedent causes (and that these partially determine our behaviour).

Though not (yet) at the same level of absurdity, the debate over the causal role of consciousness in human behaviour is being similarly muddled by worrying conceptual issues. In the case of consciousness, many assume (and sometimes define) consciousness as something non-physical and non-functional. In the case of causation, there is a lack of specificity about what exactly is doing the causing when

⁵ See, for example, Hume (1739/1975), Libet (1985), Haggard (2005), Halligan and Oakley (2021), Pockett (2004, 2006), Roediger et al. (2008), Sapolsky (2023).

appealing to conscious mental states (is it the *consciousness* of them, or the states themselves?). Furthermore, researchers often neglect (or offer implausible definitions) of what it is for something to count as a cause. The result is that psychological epiphenomenalism—the thesis that psychological evidence undermines the causal efficacy of consciousness—enjoys a readily greased slide to its desired conclusion. However, once we are armed with plausible insights about consciousness and causation, it becomes clearer how to make good sense of the causal role of consciousness in human behaviour.

3 Consciousness and Confusion

A brief note to the reader before proceeding. In this and the following section, I will be commenting on problems that are the topics of entire papers, books, and in many cases entire literatures in their own right. To justify the brevity, let me be clear that the goal is to draw attention to—and offer ways to resist and replace—unnecessarily troublesome views about consciousness and causation in the debate over psychological epiphenomenalism. Some of what follows can be read as a helpful resource for those interested in a broad overview of conceptual issues in the field. Such an overview naturally comes at the expense of doing any one topic sufficient justice, but I trust that the benefits of seeing a bit more of the forest for the trees in one place will earn this project's keep. Readers who are inclined to axe this approach can follow up on the references I will provide to the broader literature.

3.1 Is Consciousness Physical or Non-Physical?

Assertions and assumptions that consciousness is not physical abound in the scientific literature. For example, Pockett (2004) draws a distinction between “consciousness per se” and the physical neural activity that underlies it. She then goes on to express puzzlement over the idea that consciousness per se could be causally efficacious, writing: “it is difficult to see how consciousness could break into the deterministic/random chain of events in the brain and impose a new direction on these” (p. 23). In a similar vein, Pierson and Trout (2017) argue that consciousness is necessary for volitional action, but that consciousness is non-physical and distinct from “deterministic” physical brain processes (p. 67). After arguing that epiphenomenal property dualism is the logical conclusion to draw from various sources of psychological evidence on unconscious processes, Oakley and Halligan (2017) claim: “The converse question ‘How can the non-physical experiences of ‘conscious awareness’ control physical processes in the brain?’ is consequently no longer relevant” (p. 9). In a substantial literature review on the necessity of consciousness for sophisticated human action, Baumeister and colleagues (2018) admit: “We concede that it is not obvious how mental events such as conscious thoughts and feelings could influence the physical world” (p. 2).

Many do not so much argue as merely assert or assume that consciousness is not physical, and it is worth acknowledging that there is something intuitively compelling about this.⁶ After all, most of us probably have a commonsense itinerary of what it is for something to count as physical, and it can be hard to see how consciousness fits these criteria. Paradigmatically physical objects have certain properties like mass, position in space, solidity, and so on. But consciousness has none of these properties, so one might doubt that consciousness is physical.

Well, it depends on what you mean by “physical”. It turns out that defining what it is for something to be physical is much more complicated than it first appears, and there are different views on this. According to one view—the *object-based conception*—a property counts as physical if (and only if) it is the sort of property had by paradigmatically physical objects and their constituents (Stoljar 2001). On another view—the *theory-based conception*—a property counts as physical if (and only if) it is the sort of property that physical theories (like physics) tell us about (Stoljar 2001). There are many other views besides. I will not rehearse this debate here, save for noting that specifying what exactly it is for something to be physical is notoriously difficult, and so when I speak of consciousness being (or not being) physical, I will leave it open for the reader to interpret this in light of their preferred theory.⁷

What is of more general relevance here is the metaphysical doctrine of *physicalism*, which in its most basic incarnation is the thesis that everything that exists is (or bears some important relation to) the physical, and that many think must be believed to maintain any kind of scientific or rational approach to the world. One big reason for this is due to what is called the “causal closure of the physical”, which in simplified terms is the idea that every physical event has a sufficient physical cause (Kim 1998). If we accept that mental events cause physical events (e.g., consciously deciding to raise my arm causes my arm to raise), then to avoid causal overdetermination, mental events must be (or supervene on) physical events. Of course, physicalists accept that there are many things in the world that do not appear at first glance to be physical such as culture, things of a mathematical nature, morality, and indeed, consciousness. But the idea is that in the final analysis all such items will turn out to be physical or bear some important relation to it.

There are familiar arguments about consciousness being non-physical from the philosophical literature that might appear to justify assuming it is not physical. Here I will rehearse two of the most prominent among them. One is the modal argument involving thought experiments about philosophical zombies, which Chalmers (1995) formulated as follows: “A zombie is just something physically identical to me, but which has no conscious experience—all is dark inside. While this is probably empirically impossible, it certainly seems that a coherent situation is described; I can discern no contradiction in the description” (p. 96). If things like zombies are conceivable, the thought goes, then they are possible too (at least logically), which some take to show that consciousness is non-physical.

⁶ I thank an anonymous reviewer for pressing me on this point.

⁷ For an overview of this topic, see Stoljar (2010).

Jackson's (1982) knowledge argument is another powerful thought experiment against physicalism. The story is so well-known to be almost a platitude. We are asked to imagine a brilliant scientist, Mary, who has never experienced colour due to being locked up in a black and white room her entire life but has nevertheless learned everything there is to know physically about colour. Upon being let out of her room for the first time and seeing something red, we are invited to wonder whether she will learn something new. Many people have the reaction that she will learn something, namely, what it is like to see red. But she had all the physical information beforehand. Ergo, there is more to have than physical information, so physicalism is false.

Although they arrive at their conclusions through different means, the modal and knowledge arguments buttress neatly with one another to form a tandem attack on physicalism.⁸ If the conceivability and possibility of zombies goes through, then the modal argument shows that consciousness is both non-physical and epiphenomenal. For if there could be a physically and behaviourally identical being to us who lacks consciousness altogether, it is difficult to see what difference consciousness makes to our behaviour as it currently stands. If the knowledge argument goes through, then again consciousness is non-physical and epiphenomenal with respect to behaviour. Together, we have quite a potent cocktail.

How to respond? There is a substantive literature on the modal and knowledge arguments (and other prominent anti-physicalist arguments, e.g., Nagel 1974). Popular strategies for denying the modal argument include denying that we can in fact conceive of zombies in the relevant sense, while others prefer to question the link between conceivability and possibility. Strategies for denying the knowledge argument involve rejecting the claim that Mary has all the physical information, that she learns anything new upon exiting, accepting that she does learn something new but that this is capturable in physicalist terms by way of something like a new concept or ability, or by arguing that only *a priori* physicalism is undermined while *a posteriori* physicalism remains untouched. A broad overview of these (and other) issues can be found in Kind and Stoljar (2023).

One solution that is especially relevant for present purposes is the *ignorance hypothesis* (Stoljar 2006). The thought is that we are ignorant of a type of physical fact or property relevant to explaining consciousness. This ignorance makes us overconfident in thinking we can properly conceive of philosophical zombies which are supposed to be physically identical to us, or Mary who we are invited to imagine knows all the physical facts. Instead, their apparent conceivability is owed to us underestimating our own ignorance. Thus, such conceivability arguments about the incompatibility of consciousness and physicalism fail, for they presuppose knowledge of all the physical facts or properties relevant to consciousness that we cannot confidently assert we can even conceive of.

The upshot of the ignorance hypothesis is that it removes any philosophical reason for thinking that the existence of consciousness is incompatible with

⁸ Though see McGeer (2003) for a contrary view that the modal and knowledge arguments are in tension with one another.

physicalism. We do not need to put the cart before the horse and load ourselves up with the metaphysical baggage of dualism and panpsychism (see Frankish 2021), resign ourselves to the view that we do not have any feelings (eliminativism), or convince ourselves that our feelings are introspective illusions (illusionism). It also licenses us to avoid premature assumptions about the physicality or non-physicality of consciousness. As it stands, we simply do not know enough to be confidently loading the dice in either direction.

It might be thought that scientists attacking or defending psychological epiphenomenalism on empirical grounds have no obvious obligation to attend to matters in the metaphysics of mind. But things change when problematic metaphysical assumptions about the non-physicality of consciousness are being made. Given our humble epistemic state when it comes to the nature of consciousness, loading the dice against the causal efficacy of consciousness at the outset by assuming it to be non-physical is at best contingent on very controversial arguments from the metaphysics of mind, and at worst will turn out with the benefit of hindsight to be obviously false and misleading.⁹ Until conclusive evidence comes in for the physicality or non-physicality of consciousness, our best move in trying to understand how (or whether) consciousness causes behaviour is to assess the scientific evidence on impartial grounds without making anti-physicalist assumptions.¹⁰

3.2 The Functions of Consciousness

The second issue about consciousness worth addressing is targeted at those who overdraw the distinction between phenomenal and access consciousness. Ironically, the very conceptual distinction between the two was invoked to clear up a confusion about the function of consciousness (Block 1995), and yet in what follows I want to venture the suggestion that some conceptual gluing is in order. This will make the problem of the causation of behaviour by phenomenal consciousness more tractable.

Phenomenal consciousness, to reiterate, can be thought of as a generalization of the notion of feeling. Creatures or mental states are phenomenally conscious when it *feels* some way or other to be that creature or in that state. The qualitative character of our experiences—the painfulness of a kicked shin or the salty taste of Dutch liquorice—are referred to as phenomenal properties. Access consciousness, by contrast (or what is also sometimes called “psychological consciousness” or “awareness”, see Chalmers 1996), refers to mental states that can be reported on and are poised or available for rationally guiding thought and action (Block 1995). The thought is that a phenomenally conscious state of feeling a particular way need not be available for being reported or rationally guiding action, and that an access conscious state need not have a particular phenomenal feel to it.

⁹ For historical examples of structurally analogous cases against physicalism involving ignorance (and where the ignorance about the relevant physical facts was cleared up by a gain in knowledge), see Stoljar (2006, pp. 123–141).

¹⁰ I am indebted to an anonymous reviewer for suggesting this impartial view (as opposed to taking physicalism as the default).

To adapt an example from Block (1995), suppose I am sitting at my desk typing up this paper and a leaf-blower is going outside my window (as in fact it has been—*all morning*). I was focused on the paper, but then I paused, raised my head, and suddenly realized that there was—and had been for some time—a leaf blower gurgling away. Given that the leaf-blower was going that whole time, it is plausible to think that I was conscious of it in the phenomenal sense—there was something it was like, even if only minimally, for its constant hum to be a part of my ongoing phenomenal experience. But I was also unconscious of it in the access sense because of how focused I was on my paper. I had no cognitive access to it until now, at which point it is both phenomenally and access conscious.¹¹

The phenomenal/access distinction has undoubtedly been a valuable tool in contributing to our ability to conceptually manipulate and get a deeper handle on the functions of consciousness. One prominent reason is that researchers are apt to run the two together, claiming to have explained the functions of phenomenal consciousness when only offering an account of access or awareness (Block 1995; Chalmers 1996). However, trouble arises when this distinction is overdrawn. By “overdrawn” I mean that researchers think about access consciousness as the exclusively functional kind of consciousness, whereas phenomenal consciousness is left by the functional wayside. Defining access consciousness in terms of functional roles that consciousness serves, and phenomenal consciousness solely as what it feels like to experience mental states, leaves the latter vulnerable to being presented as a kind of icing that can be cleanly scraped off the behavioural cake. To borrow a delightful label from Frankish (2021), this is to “depsychologize” phenomenal consciousness, and the problems that ensue—including having to contend with epiphenomenalism—are the well-deserved rewards for this fragmented thinking.

In practice, defining phenomenal consciousness as something beyond the performance of functions greases the slide to accepting thought experiments like zombies and empirical studies from psychology that claim it makes no difference to our behaviour. But just as we should give pause before accepting that science has shown we do not have free will if “free will” means something questionable, a similar pause is warranted before too readily accepting the idea that we can always cleave phenomenal consciousness asunder from the performance of functions. I say “always” because I do not want to rule out the possibility that some phenomenal states might be functionless some of the time. What I object to is their blanket categorization of being so—and this is what you get by overdrawing the phenomenal/access distinction.

Take the experience of pain and pleasure as examples. Feeling pain and pleasure are states that are paradigmatically phenomenally conscious (i.e., there is something painful or pleasurable it feels like for the subject to be in them) and have a typical profile of causes and effects. In the absence of interfering motives, experiencing pain and pleasure causes us to behave in certain ways, namely, to avoid pain and pursue pleasure, due to their phenomenal character or way they feel (Morch 2018).

¹¹ For arguments drawing on empirical work to demonstrate that phenomenal consciousness overflows cognitive access, see Block (2007).

But such phenomenally conscious states might also vary in the degree to which they are access conscious too. As these feelings enjoy a higher degree of access consciousness, then we can expect the functional role of those feelings to be enhanced in various ways. The feeling might have a wider profile of causes and effects than it otherwise would have, the feeling might make the avoidant or seeking behaviour more likely to occur, or the feeling might generate higher attentional involvement in the feeling or the stimulus that evoked it.

There are familiar arguments that try to show that phenomenal consciousness comes apart from functions (Block 1995, 2007; Chalmers 1996), but equally there are plenty of researchers who convincingly challenge this dissociation in careful detail (Carruthers 2017; Cleeremans and Tallon-Baudry 2022; Cohen & Dennett, 2011; Frankish 2021; for a recent overview, see Niikawa et al. 2022). I will not repeat those arguments here, save for emphasizing that there need not be anything about the phenomenal/access distinction *per se* that precludes the phenomenal character of various mental states from being constitutive of their functional role.¹² Indeed, if we think of access consciousness as the *enhancement* of a state's functional role, then this is perfectly consistent with phenomenal states having functional roles too—ones that can be enhanced when also conscious in the access sense.

The bottom line is that defining or assuming phenomenal consciousness can always be separated from functions is—much like the oft-defined or assumed non-physicality of consciousness—a controversial thesis that is apt to muddy the water with respect to the debate on psychological epiphenomenalism. Just as it would be a mistake to unfairly stack the deck in favour of epiphenomenalism by assuming that consciousness is non-physical, so too it is a mistake to assume that phenomenal consciousness is always divorceable from the performance of functions.

4 Causation and Confusion

Having dealt with these two issues about consciousness that generate confusion about how it could make a difference to our behaviour, let me move on to some further issues that generate confusion in the consciousness literature, this time about causation. First, researchers often talk of the causal efficacy (or lack thereof) of “conscious thoughts”, “conscious decisions”, “conscious intentions”, and so forth, as if this is not ambiguous between the *consciousness* of the mental states doing the causal work on the one hand, and the *mental states* (of which we just happen to be conscious) being causally efficacious on the other. The empirical evidence adduced for and against psychological epiphenomenalism often runs afoul of this distinction. Second, what exactly is meant by “causation” is often suboptimal or undefined in the consciousness literature, and although this is a somewhat more minor point, I think addressing it will help to avoid confusion in future work.

¹² I am grateful for discussions with Daniel Stoljar on this point.

4.1 What is Doing the Causing?

The first source of confusion concerns what precisely is doing the causing when researchers refer to conscious mental states like “conscious thoughts”, “conscious intentions”, “conscious decisions”, and variations thereof. To ask whether a conscious mental state is among the causes of our actions admits of at least two distinct interpretations of what is doing the causal work. On the one hand, there is the question of whether a *mental state* (of which we just happen to be conscious) is doing the causing. On the other hand, there is the question of whether the fact that the mental state is *conscious* is doing the causing. Compare the following claims:

CONSCIOUS INTENTION₁ I have a conscious intention to perform an action, and the action is caused by my *intention* to perform it.

CONSCIOUS INTENTION₂ I have a conscious intention to perform an action, and the action is caused by my *consciousness* of my intention to perform it.

Although these are superficially similar theses, they are in fact substantially distinct (even though they may frequently co-occur in practice). Think of it this way. Intentions are mental states and paradigmatic causes of action. We know that intentions can be unconscious in all the relevant senses (e.g., phenomenal and access) and still cause behaviour (Mele 2009). So, an intention’s being a *conscious* intention is not necessary for it to influence our behaviour, given that it can exert a causal influence in virtue of the kind of mental state it is. What this means is that saying that a “conscious intention” caused an action leaves it open as to what work (if any) is being done by the *consciousness* of the intention in addition (or as opposed to) to the intention itself. A similar point applies to mental states more broadly.

With this distinction on the table, my worry is that various attempts to marshal arguments and evidence for and against psychological epiphenomenalism often fall afoul of it. What starts out as a promise to attack or defend the causal efficacy of consciousness in the introductory sections of various papers ends up slipping into a debate over the causal efficacy of mental states like thoughts, intentions, beliefs, desires, and so on, and one is left feeling like the victim of a bait-and-switch.¹³

Before I discuss some examples of what I have in mind, it is important to note that one of the standard tools that scientists use to test whether consciousness causes behaviour is an experimental design. The idea is that consciousness is manipulated as the independent variable between a control group (with no manipulation) and the experimental group (with the manipulation), and the effects on behaviour are measured as the dependent variable. If everything else in the experiment is held fixed,

¹³ For examples of researchers on both sides of the debate who sometimes fall afoul of this distinction, see Bargh (2013), Baumeister and Masicampo (2010), Baumeister et al. (2011, 2018), Halligan and Oakley (2021), Lieberman (2009), Masicampo and Baumeister (2013), Roediger et al. (2008), Sklar et al. (2021).

any statistically significant effects are taken as evidence that consciousness was among the causes of behaviour.

This would be perfectly acceptable if consciousness were in fact the independent variable, but this is not always the case. One of the more influential articles that has pushed back against psychological epiphenomenalism is Baumeister, Masicampo, and Vohs's (2011) comprehensive review titled "Do Conscious Thoughts Cause Behaviour?".¹⁴ Throughout the paper, they marshal an impressive array of evidence in support of the role that consciousness plays in human behaviour, concluding that: "The evidence for conscious causation of behavior is profound, extensive, adaptive, multifaceted, and empirically strong" (p. 351). In what follows, I will briefly summarize the first two categories of evidence they appeal to: **Mental Simulation/Practice** and **Planning, Intending, Anticipating**. This is obviously not an exhaustive list, but it is enough to get a feel for the issue I want to draw attention to.

Mental Simulation/Practice Imagining oneself doing something can increase the likelihood or efficacy of doing it. Gregory et al. (1982) found that participants who were given information about a cable television service were more likely several months later to be subscribed to the service if they were in the group that were instructed to imagine themselves getting and using cable television. Imagining oneself voting can increase the probability of voting in an upcoming election (Libby et al. 2007). Imagining oneself performing an athletic, artistic, or physical skill effectively can improve performance across a range of different activities. In the case of golf, pairing mental practice with physical practice improved performance compared to physical practice alone (Brouziyne and Molinaro 2005).

Planning, Intending, Anticipating Making specific goal-oriented plans in the form of "implementation intentions", and anticipating emotional states, both have significant effects on behaviour. Implementation intentions promote the attainment of vague goal intentions ("I intend to x") by specifying when, where, and how we will perform the intended goal ("when situation x arises, I will perform response y"). An impressive amount of evidence demonstrates that furnishing vague intentions, desires, and goals with specific implementation intentions makes us far more likely to engage in the desired behaviour, sometimes doubling or tripling success rates (Gollwitzer 1999; Gollwitzer and Sheeran 2006). Studies have also found that anticipated emotions—especially regret and guilt—motivates people to behave in ways that aim to alleviate the emotions (Lindsey 2005; Risen and Gilovich 2007).

These are all very interesting findings, but what do they tell us about consciousness? In the case of mental simulation and practice, one can reasonably conclude that different kinds of imaginings can have different effects on behaviour. In the case

¹⁴ I take it that Baumeister, Masicampo, and Vohs's (2011) have something akin to *access consciousness* in mind. They explicitly set aside the phenomenal variety in the definition section, and shortly after claim that reportable inner states constitute the usual criterion for determining whether something is conscious (see pp. 333–334).

of planning, intending, and anticipating, it is clear that forming very specific plans in the form of implementation intentions can make it much more likely that our intentions will be effective in action. We can also conclude that being in or anticipating various emotional states can change how we act. Notice, however, that all this shows us is that various mental states, events, or processes cause behaviour. While these were all conscious in various senses in the above experiments, consciousness was not the independent variable, and so causation by consciousness cannot be inferred. To say that a conscious imagining, conscious intention, or conscious emotion caused behaviour does not tell us that any work is done by their being *conscious* imaginings, intentions, and emotions.¹⁵

Of course, arguments can be given for why the consciousness of these mental states do feature in the causes of action.¹⁶ But the point is that more does need to be said to establish this conclusion. Let it also be noted that there are plenty of different sources of evidence that researchers appeal to in defending the causal efficacy of consciousness, and not all of it runs afoul of the distinction I have raised here.¹⁷ In any case, the takeaway is that conscious mental states being among the causes of behaviour does not—without a further step or premise—establish that the consciousness of those mental states numbers among the causes too. As I see it, being more precise and careful on this point will be a helpful contribution to further research in the debate on psychological epiphenomenalism.

4.2 Defining Causation

When making claims about whether consciousness causes behaviour, it is important to be clear about what it means for something to count as a cause. However, causation is often not defined at all in the literature about the causal efficacy of consciousness. In a sustained argument showing why Libet-style experiments do not show that consciousness is epiphenomenal, Stockdale (2022) argues that consciousness is among the causes of our actions without defining what it is for something to be a cause. Baumeister, Masicampo, and Vohs's (2011) comprehensive literature view similarly offers no definition of causation. Oakley and Halligan (2017), Halligan and Oakley (2021), and Pockett (2004) all argue that the psychological evidence on unconscious processes shows that consciousness is a causally inert epiphenomenon, despite never once going into sufficient detail about what they mean by causation. Note that all these authors—and many others besides—make use of “cause”,

¹⁵ It might be thought that one or more of these are necessarily conscious. However, it is widely accepted that intentions and emotions can be unconscious (and it is at least entertained that implementation intentions could be unconscious, see Mathieson 2024, pp. 138–139; Mele 2009, pp. 131–144). It is less obvious that imagination can be unconscious, but there are plausible arguments for this view (for a critical overview, see Kind 2021).

¹⁶ For arguments to this effect in the context of implementation intentions, see Mathieson (2024, pp. 134–135) and Mele (2009, pp. 131–144).

¹⁷ Thanks to an anonymous reviewer for pointing this out. See, for example, Baumeister and Masicampo (2010), who argue that consciousness performs a variety of social and interpersonal functions that cannot be achieved unconsciously.

“causation”, and their derivatives frequently. The problem is that it is never quite made explicit what exactly a cause *is*.

It is one thing to note that researchers are failing to define key terms. But it is another thing to say why this matters. After all, it might be argued that it is perfectly appropriate for researchers to avoid taking a stand on how causation should be understood lest they unnecessarily alienate readers. Think back to the point raised earlier about the difficulty in defining “physical”, and the recommendation to leave things open so that readers can interpret the arguments according to their preferred view. Researchers might think something similar holds in the case of causation too, in that their arguments go through on various conceptions of causation.¹⁸

Two points can be made in response. First, a great deal of talking past each other, confusion, conflation, and misunderstanding occurs when people employ terms that have multiple meanings in an imprecise way. Think about how much consciousness research has benefitted from its authors carefully defining key terms like “consciousness”. Nobody would dream of writing a paper on consciousness anymore without specifying what kind of consciousness they have in mind, and this is generally agreed to be good practice. Whether or not you agree with a given author’s take on consciousness, it can hardly be denied that it is at least helpful to have a clear idea of what they have in mind. In my view, it would likewise be good practice going forward to apply the same principle in the case of what one has in mind by causation.

Second, I am sceptical as to whether the absence of clear definitions of causation in the literature on epiphenomenalism really is analogous to the move about the nature of the physical. For one thing, the persistence of some very confused reasoning about causation in the scientific literature reveals a far less optimistic situation than this objection suggests. One useful corollary of having to think carefully about what it is for something to count as a cause is that mistakes about causation are much less likely to occur. This supplies a further reason to encourage researchers to get clearer on what they mean by consciousness causing behaviour.

Take Libet’s (1985) famous experiments in which the consciousness of intentions to perform simple actions were found to occur some 350ms after the onset of neural activity preparing the action, but 200ms or so before the action itself. Commenting on this work, Roediger et al. (2008) write: “Clearly, conscious intention cannot cause an action if a neural event that precedes and correlates with the action comes before conscious intention” (p. 208). Haggard (2005) similarly writes: “The seminal studies of Benjamin Libet...suggested that conscious intention occurs after the onset of preparatory brain activity. It cannot therefore cause our actions, as a cause cannot occur after its effect” (p. 291).

The above passages evince precisely the kind of confused reasoning about causation that getting clearer about what it means for something to count as a cause would likely have ameliorated. The fact that neural events precede the consciousness of our intentions to act in no way shows that consciousness cannot cause actions. Actions are processes (or events—I make no distinction here), processes have parts, and consciousness can certainly still be among the causes of the action despite being

¹⁸ Thanks to an anonymous reviewer for raising this objection.

preceded by neural events. As for causes not being able to occur after their effects, it is important to clarify that Haggard is not claiming that the consciousness of the intention occurs after the action. Like Roediger and colleagues, the reasoning seems to be that consciousness being caused by prior neural activity makes it merely an effect rather than another possible cause of the action. But the error is assuming that for something to count as a cause it cannot have been caused by other things. This would involve a very odd break in reality—a break that no view of causation is committed to. Causes have causes, and consciousness is no exception.¹⁹

So far, I have been trying to motivate the idea that it would improve the epiphenomenalist debate to be clearer on what is meant by causation. But it would be remiss of me not to acknowledge some progress on this front. For example, Baumeister et al.'s (2018) recent response to psychological epiphenomenalism involves an appeal to Mackie's (1980) INUS model of causation, where a cause is defined as an Insufficient but Necessary part of an Unnecessary but Sufficient Condition. According to this view, consciousness qualifies as a cause of behaviour if it is a necessary condition or part of the processes that led to the behaviour. While this is a step in the right direction, there are better theories of causation to hinge one's arguments upon—especially given that Mackie's INUS model is largely unpopular with contemporary philosophers due to being susceptible to notorious counterexamples.²⁰

Happily, more plausible accounts of causation are at hand. When philosophers and psychologists ask whether consciousness causes behaviour, in many cases what they are interested in is whether consciousness *makes a difference* to what we do such that, if consciousness were absent, behaviour would (or would not) be the same. Understanding causation according to a “difference-making” account is well suited for this purpose, being labelled as the most natural account to employ in the context of causal attribution and explanation in the sciences, and in particular for mental causation (List and Menzies 2017). According to this view, causation is a form of counterfactual or probabilistic dependence, which List and Menzies (2017) express counterfactually as follows:

Positive Conditional If *C* were to occur, then *E* would occur.

Negative Conditional If *C* were not to occur, then *E* would not occur.

However, it does make a difference (not to be a pun) whether the difference-making account is stated counterfactually or probabilistically, as according to some philosophers there would hardly ever be causation on the counterfactual account (Hajek 2021). One problem is that the “would” in the consequent is too strong. *C* occurring

¹⁹ For a thorough refutation of several examples of researchers engaging in similarly confused reasoning about causation, see Mele (2009, pp. 70–73).

²⁰ Some acknowledged by Mackie (1980) himself, such as the Manchester factory hooters case: “The sounding of factory hooters in Manchester [at 5pm] may regularly be followed by, but does not cause, London workers leaving their work” (p. 81). The regularity and the temporal order of the sound makes it an INUS condition, but it does not follow that the sounding of the alarm causes London workers to leave their work.

does not necessitate *E* occurring, and vice versa. Instead, we might think that *C* occurring increases the conditional probability of *E* occurring, and vice versa.

It is beyond the scope of this paper to offer a fully developed and defended view of causation, but the point is that the difference-making account is one attractive way to understand causation in the context of psychological epiphenomenalism.²¹ Moreover, accounts that argue for or against psychological epiphenomenalism often implicitly operate with something approximating a difference-making account of causation, and making this implicit assumption explicit would help to bolster the clarity of their views. The general lesson here is that being clear and explicit about what one means by causation, and being more careful in what view of causation one adopts, would improve existing accounts and help to avoid implausible views about the nature of causation in the debate on psychological epiphenomenalism.

5 Conclusion

Let us take stock. Issues about the oft-assumed non-physicality and non-functionality of consciousness were identified and addressed. Ambiguity in what is doing the causing, and either implausibly defining causation or not defining it at all, were also discussed. These issues—sometimes individually, sometimes all together—contribute to unnecessary scepticism and confusion about how consciousness features among the causes of human behaviour. Given that critics and defenders of psychological epiphenomenalism are largely in the business of assessing whether psychological evidence supports the claim that consciousness is causally inert, weaving in these sorts of assumptions is apt to unnecessarily complicate this task. It cannot be repeated often enough that such assumptions also conveniently happen to grease the slide to epiphenomenalism.

It goes without saying that there are plausibly many other sources of confusion about the nature of consciousness and causation which remain to be addressed. I have tried to express the peculiar issues that I have found especially confusing as best as I am able, and to suggest a clearer path going forward. Without concerted efforts at conceptual hygiene in this fantastically interesting (and at times exasperatingly unwieldy) literature, we are apt to become unwilling participants in the generation of confusion about how consciousness causes behaviour. Much unwarranted scepticism occurs as a result, and much is to be gained by clearing it up.

²¹ As an anonymous reviewer has pointed out, it is important to note that the difference-making account is—much like any account of causation—subject to various objections. One specific issue worth mentioning for difference-making accounts of causation is the “fat-handedness” of many independent variables (Eronen 2020). The worry here is that certain variables—and consciousness might be among them—are extremely hard to intervene on in an isolated way without also changing other related variables that may also influence the outcome. An overview of this and other issues about causal inference in human behaviour (along with various suggestions about how to overcome them) can be found in Bailey et al. (2024). See also the collection of essays in Anjum and Mumford (2018) and Beebe et al. (2017) for further theorizing about the nature of causation in general, and difference-making in particular. Hu (forthcoming) also offers some insightful work on this topic, particularly the exhortation to avoid translating metaphysical oversights about causation into methodological ones.

Acknowledgements Thanks to Brian Garrett, Alan Hajek, Frank Jackson, Andrew Lee, Jonne Maas, Tori McGeer, and Daniel Stoljar for comments and discussion on the material in previous versions of this paper. I also benefitted tremendously from discussions on interventionist accounts of causation in the Values in Science reading group at the department of History and Philosophy of Science, University of Cambridge. I am further grateful to audience members at the Philosophy of Mind work in progress group at ANU where an early version of this paper was presented, and to two extremely helpful and exacting referees whose critical feedback improved the paper significantly.

Funding This research is supported by an Australian Government Research Training Program (RTP) Scholarship.

Declarations

Competing Interests The author declares no competing interests.

Conflict of Interest None.

References

- Anjum, Rani, and Stephen Mumford. 2018. *Causation in science and the methods of scientific discovery*. Oxford: Oxford University Press.
- Balaguer, Mark. 2019. Free will, determinism, and epiphenomenalism. *Frontiers in Psychology* 9: 1–14. <https://doi.org/10.3389/fpsyg.2018.02623>.
- Bargh, John. 2013. Social psychology cares about causal conscious thought, not free will per se. *British Journal of Social Psychology* 52: 228–230.
- Bargh, John. 2017. *Before you know it: he unconscious reasons we do what we do*. New York: Touchstone.
- Bargh, John, and Ezequiel Morsella. 2008. The unconscious mind. *Perspectives on Psychological Science* 3: 73–79.
- Bargh, John, and Erin Williams. 2006. The automaticity of social life. *Current Directions in Psychological Science* 15: 1–4.
- Baumeister, Roy, and E. Masicampo. 2010. Conscious thought is for facilitating social and cultural interactions: How mental simulations serve the animal-culture interface. *Psychological Review* 117: 945–971.
- Baumeister, Roy, Stephen Lau, Heather Maranges, and Cory Clark. 2018. On the necessity of consciousness for sophisticated human action. *Frontiers in Psychology* 9: 1–6.
- Baumeister, Roy, E. Masicampo, and Kathleen Vohs. 2011. Do conscious thoughts cause behaviour? *Annual Review of Psychology* 62: 331–361.
- Beebe, Helen, Christopher Hitchcock, and Huw Price. 2017. *Making a difference: Essays on the philosophy of causation*. Oxford: Oxford University Press.
- Block, Ned. 1995. On a confusion about a function of consciousness. *Behavioural and Brain Sciences* 18: 227–247.
- Block, Ned. 2007. Consciousness, accessibility, and the mesh between psychology and neuroscience. *Behavioural and Brain Sciences* 30: 481–499.
- Brouziyne, Majid, and C. Molinaro. 2005. Mental imagery combined with physical practice of approach shots for golf beginners. *Perceptual and Motor Skills* 101: 203–211.
- Carruthers, Peter. 2017. Block's overflow argument. *Pacific Philosophical Quarterly* 98: 65–70.
- Chalmers, David. 1995. Facing up to the problem of consciousness. *Journal of Consciousness Studies* 2: 200–219.
- Chalmers, David. 1996. *The conscious mind: In search of a fundamental theory*. New York: Oxford University Press.
- Cleeremans, Axel, and Catherine Tallon-Baudry. 2022. Consciousness matters: Phenomenal experience has functional value. *Neuroscience of Consciousness*. <https://doi.org/10.1093/nc/niac007>.
- Cohen, Michael, and Daniel Dennett. 2011. Consciousness cannot be separated from function. *Trends in Cognitive Sciences* 15: 358–364.

- Dijksterhuis, Ap, and Loran Nordgren. 2006. A theory of unconscious thought. *Perspectives on Psychological Science* 1: 95–109.
- Drew, Bailey, Alexander Jung, Adriene Beltz, Markus Eronen, Christian Gische, Ellen Hamaker, Konrad Kording, Catherine Lebel, Martin Lindquist, Julia Moeller, Adeel Razi, Julia Rohrer, Baobao Zhang, and Kou Murayama. 2024. Causal inference on human behaviour. *Nature Human Behaviour* 8: 1448–1459.
- Eronen, Markus. 2020. Causal discovery and the problem of psychological intervention. *New Ideas in Psychology* 59: 1–8. <https://doi.org/10.1016/j.newideapsych.2020.100785>.
- Frankish, Keith. 2021. Panpsychism and the depsychologization of consciousness. *Aristotelian Society Supplementary Volume* 95: 51–70.
- Goldstein, Ariel, and Ran Hassin. 2017. Commentary: Definitely maybe: Can unconscious processes perform the same functions as conscious processes? *Frontiers in Psychology* 8: 1–4.
- Gollwitzer, Peter. 1999. Implementation intentions: Strong effects of simple plans. *American Psychologist* 54: 493–503.
- Gollwitzer, Peter, and Paschal Sheeran. 2006. Implementation intentions and goal achievement: A meta-analysis of effects and processes. *Advances in Experimental Social Psychology* 38: 69–119.
- Gregory, W, Robert Cialdini, and Kathleen Carpenter. 1982. Self-relevant scenarios as mediators of likelihood estimates and compliance: Does imagining make it so? *Journal of Personality and Social Psychology* 43: 89–99. <https://doi.org/10.1037/0022-3514.43.1.89>
- Haggard, Patrick. 2005. Conscious intention and motor cognition. *Trends in Cognitive Sciences* 9: 290–295.
- Hajek, Alan. 2021. Counterfactual scepticism and antecedent contextualism. *Synthese* 199: 637–659.
- Halligan, Peter, and David Oakley. 2021. Giving up on consciousness as the ghost in the machine. *Frontiers in Psychology* 12: 1–16. <https://doi.org/10.3389/fpsyg.2021.571460>.
- Hassin, Ran. 2013. Yes it can: On the functional abilities of the human unconscious. *Perspectives on Psychological Science* 8: 195–207.
- Hesselmann, Guido, and Pieter Moors. 2015. Definitely, maybe: Can unconscious processes perform the same functions as conscious processes? *Frontiers in Psychology* 6. <https://doi.org/10.3389/fpsyg.2015.00584>
- Hu, Lily. Forthcoming. Normative facts and causal structure. *Journal of Philosophy*.
- Hume, David. 1739/1975. In *A treatise of human nature*. Oxford: Clarendon.
- Jackson, Frank. 1982. Epiphenomenal Qualia. *The Philosophical Quarterly* 32: 127–136.
- Kim, Jaegwon. 1998. *Mind in a physical world*. Cambridge, MA: MIT Press.
- Kind, Amy. 2021. Can imagination be unconscious? *Synthese* 199: 13121–13141.
- Kind, Amy, and Daniel Stoljar. 2023. *What is consciousness? A debate*. New York: Routledge.
- Lavazza, Andrea. 2019. Why cognitive sciences do not prove that free will is an epiphenomenon. *Frontiers in Psychology* 10: 1–11. <https://doi.org/10.3389/fpsyg.2019.00326>.
- Libby, Lisa, Eric Shaeffer, Richard Eibach, and Jonathan Slemmer. 2007. Picture yourself at the polls: Visual perspective in mental imagery affects self-perception and behavior. *Psychological Science* 18: 199–203.
- Libet, Benjamin. 1985. Unconscious cerebral initiative and the role of conscious will in voluntary action. *Behavioural and Brain Sciences* 8: 529–566.
- Lieberman, Matthew. 2009. What zombies can't do: A social cognitive neuroscience approach to the irreducibility of reflective consciousness. In *Two minds: Dual processes and Beyond*, eds. Jonathan Evans and Keith Frankish, 293–316. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199230167.003.0013>
- Lindsey, Lisa. 2005. Anticipated guilt as behavioral motivation: An examination of appeals to help unknown others through bone marrow donation. *Human Communication Research* 31: 453–481.
- List, Christian, and Peter Menzies. 2017. My brain made me do it: The exclusion argument against free will, and what's wrong with it. In *Making a Difference: Essays on the Philosophy of Causation*, eds. Helen Beebe, Christopher Hitchcock, and Huw Price, 269–285. Oxford: Oxford University Press. <https://doi.org/10.1093/oso/9780198746911.003.0014>
- Mackie, John. 1980. *The cement of the universe*. Oxford: Oxford University Press.
- Masicampo, E, and Roy Baumeister. 2013. Conscious thought does not guide moment-to-moment actions—it serves social and cultural functions. *Frontiers in Psychology* 4: 1–5. <https://doi.org/10.3389/fpsyg.2013.00478>.
- Mathieson, Darryl. 2024. Psychological epiphenomenalism. *Journal of Consciousness Studies* 31: 120–143.

- McGeer, Victoria. 2003. The trouble with Mary. *Pacific Philosophical Quarterly* 84: 384–393.
- Mele, Alfred. 2009. *Effective intentions: The power of conscious will*. New York: Oxford University Press.
- Mele, Alfred. 2018. Free will, moral responsibility, and scientific epiphenomenalism. *Frontiers in Psychology* 9: 1–8. <https://doi.org/10.3389/fpsyg.2018.02536>.
- Montague, Pendleton. 2008. Free will. *Current Biology* 18: 584–585.
- Morch, Hedda Hassel. 2018. The evolutionary argument for phenomenal powers. *Philosophical Perspectives* 1: 293–316.
- Nagel, Thomas. 1974. What is it like to be a bat? *The Philosophical Review* 83: 435–450.
- Nahmias, Eddy. 2002. When consciousness matters: A critical review of Daniel Wegner's the illusion of conscious will. *Philosophical Psychology* 15: 527–554.
- Nahmias, Eddy. 2011. Intuitions about free will, determinism, and bypassing. In *The Oxford Handbook of Free Will: Second Edition*, ed. Robert Kane, 555–577. Oxford: Oxford University Press.
- Niikawa, Takuya, Katsunori Miyahara, Hiro Hamada, and Satoshi Nishida. 2022. Functions of consciousness: Conceptual clarification. *Neuroscience of Consciousness* 1: 1–10. <https://doi.org/10.1093/nc/niac006>.
- Oakley, David, and Peter Halligan. 2017. Chasing the rainbow: The non-conscious nature of being. *Frontiers in Psychology* 8: 1–16. <https://doi.org/10.3389/fpsyg.2017.01924>.
- Pierson, Lee, and Monroe Trout. 2017. What is consciousness for? *New Ideas in Psychology* 47: 62–71.
- Pockett, Susan. 2004. Does consciousness cause behaviour? *Journal of Consciousness Studies* 11: 23–40.
- Pockett, Susan. 2006. The neuroscience of movement. In *Does Consciousness Cause Behaviour? An Investigation of the Nature of Volition*, eds. Susan Pockett, William Banks, and Shaun Gallagher, 9–24. Cambridge, MA: MIT Press.
- Reber, Arthur, and Rhianon Allen. 2022. *The cognitive unconscious: The first half-century*. Oxford: Oxford University Press. <https://doi.org/10.1093/oso/9780197501573.001.0001>
- Risen, Jane, and Thomas Gilovich. 2007. Another look at why people are reluctant to exchange lottery tickets. *Journal of Personality and Social Psychology* 93: 12–22.
- Roediger, Henry, Michael Goode, and Franklin Zaromb. 2008. Free will and the control of action. In *Are we free? Psychology and Free Will*, eds. John Baer, James Kaufman, and Roy Baumeister, 205–225. New York: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195189636.003.0010>
- Sapolsky, Robert. 2023. *Determined: A science of life without free will*. New York: Penguin.
- Sklar, Asael, and Ran Hassin. 2014. The human unconscious: A functional perspective. In *Dual-Process Theories of the Social Mind*, eds. Jeffrey Sherman, Bertram Gawronski, and Yaacov Trope, 299–313. The Guilford Press. <https://doi.org/10.31234/osf.io/evcgry>
- Sklar, Asael, Rasha Kardosh, and Ran Hassin. 2021. From non-conscious processing to conscious events and back: A minimalist approach. *Neuroscience of Consciousness* 7: 1–18.
- Soon, Chun, Marcel Brass, Hans-Jochen Heinze, and John-Dylan Haynes. 2008. Unconscious determinants of free decisions in the brain. *Nature Neuroscience* 11: 543–545.
- Stockdale, Bradford. 2022. The Libet paradigm and a dilemma for epiphenomenalism. *Philosophical Psychology*. <https://doi.org/10.1080/09515089.2022.2130744>.
- Stoljar, Daniel. 2001. Two conceptions of the physical. *Philosophy and Phenomenological Research* 62: 253–281.
- Stoljar, Daniel. 2006. *Ignorance and imagination: The epistemic origin of the problem of consciousness*. New York: Oxford University Press.
- Stoljar, Daniel. 2010. *Physicalism*. New York: Routledge.
- Wilson, Timothy. 2002. *Strangers to ourselves: Discovering the adaptive unconscious*. Cambridge, MA: Harvard University Press.
- Wegner, Daniel. 2002. *The illusion of conscious will*. MIT Press.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.