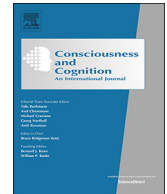




Contents lists available at ScienceDirect

Consciousness and Cognition

journal homepage: www.elsevier.com/locate/concog

The global workspace theory, the phenomenal concept strategy, and the distribution of consciousness



Dylan Black

Xi'an Jiaotong University, 28 Xianning W Rd, Jiao Da Shang Ye Jie Qu, Beilin Qu, Xian Shi, Shaanxi Sheng, China

ARTICLE INFO

Keywords:

Global workspace
Phenomenal concept strategy
Explanatory gap
Functions of consciousness
Comparative psychology

ABSTRACT

Peter Carruthers argues that the global workspace theory implies there are no facts of the matter about animal consciousness. The argument is easily extended to other cognitive theories of consciousness, posing a general problem for consciousness studies. But the argument proves too much, for it also implies that there are no facts of the matter about human consciousness. A key assumption is that scientific theories of consciousness must explain away the explanatory gap. I criticize this assumption and point to an alternative strategy for defending scientific theories of consciousness, one that better reflects the ongoing scientific practice. I argue there are introspectable inferential connections from phenomenal concepts to functional concepts that scientists can use to individuate the global workspace in terms of capacities that animals and humans share.

Carruthers (2018, 2019) argues that the global workspace theory of consciousness implies there are no facts of the matter about animal consciousness. Although animals share cognitive architecture that is similar in various ways to the global workspace, there are no facts of the matter about what kind of similarity would suffice for consciousness. If he is right, comparative psychologists should stop asking questions about animal consciousness.

In what follows I argue that Carruthers' argument should be rejected because it implies, not only that there are no facts of the matter about animal consciousness, but also that there are no facts of the matter about human consciousness. To address Carruthers' argument, I offer an alternative way to defend scientific theories of consciousness in the face of the explanatory gap that is consistent with facts of the matter about animal consciousness. In particular, I challenge the assumption that the chief *explanandum* for scientific theories of consciousness is the inferential independence of phenomenal concepts from physical/functional concepts. Scientific theories of consciousness rarely aim to explain away the explanatory gap. Rather, they aim to reveal the physical/functional basis of consciousness by finding its neural correlates. Philosophers have long argued that correlation alone cannot prove identity in the face of the explanatory gap, but I argue that there are inferential connections between phenomenal concepts and physical/functional concepts that make theorizing about the physical/functional basis of consciousness not so different from theorizing about the unobservable basis of other natural phenomena. If so, Carruthers' argument can be rejected, along with other objections to the science of consciousness. The broader goal of this paper is to make philosophical sense out of the ongoing scientific practice.

1. The phenomenal concept strategy and the global workspace theory

Most scientific theories of consciousness are supported by evidence of correlation between phenomenal states and physical/functional states of the nervous system. But philosophers argue that, in the face of the explanatory gap, correlation is not enough to

E-mail address: dylan.jacob.black@gmail.com.

<https://doi.org/10.1016/j.concog.2020.102992>

Received 15 September 2019; Received in revised form 27 May 2020; Accepted 13 July 2020
1053-8100/© 2020 Elsevier Inc. All rights reserved.

infer identity.¹ While this poses no difficulty for the minority of scientists who endorse property dualism,² it remains a challenge for the majority of scientists who are both realists and physicalists.³ Presently, the only mainstream response to the explanatory gap that is consistent with both physicalism and realism is the phenomenal concept strategy.⁴ According to the strategy, the reason phenomenal states resist explanation in terms of physical/functional processes is that phenomenal concepts are inferentially independent of physical/function concepts.⁵ In other words, correct applications of phenomenal concepts do not entail applications of physical/function concepts—and vice versa. Relatedly, phenomenal facts do not entail non-trivial physical/functional facts, and vice versa. Brian Loar, who largely originated the phenomenal concept strategy, argues that phenomenal concepts are a species of recognitional concepts because they rely on a disposition to judge that certain introspectable properties belong to the same type.⁶ Phenomenal concepts refer to the same physical/functional properties as certain directly referring physical/functional concepts, even though neither sort of concept can be inferentially derived from the other. For several decades, phenomenal concepts have been the subject of much debate in the philosophy of mind. For the most part, I will avoid these debates. All that matters for present purposes is that phenomenal truths are inferentially independent of physical/functional truths.⁷

While the phenomenal concept strategy is arguably the only response to the explanatory gap that has the potential to be consistent with the science of consciousness, many of its advocates argue that phenomenal concepts may be too opaque to support a legitimate science of consciousness.⁸ Philosopher and cognitive scientist Peter Carruthers, however, argues that the phenomenal concept strategy and the science of consciousness can supplement one another. As he sees the matter, philosophy is needed to explain why the gap arises, and science is needed to explain the introspective and higher-order capacities that explain inferential independence.⁹

In Carruthers' view, the global workspace theory of consciousness is in the best position to explain inferential independence.¹⁰ Originally set forth by cognitive psychologist Bernard Baars, the theory proposes that the distinction between conscious mental states and unconscious mental states is best explained by the accessibility of their representational content.¹¹ Whereas the contents of unconscious mental states are restricted to cognitive systems that perform relatively automatic, domain-specific tasks, the contents of conscious mental states are made widely available to many different cognitive systems via the global workspace. The content of the global workspace is decided by a competition for attention between many different unconscious processes all running in parallel. The winners of this competition have their contents consolidated into a unified scene that informs decision processes in real-time. Their contents are “globally broadcasted” to the executive processes responsible for decision making and guided control over action, along with many other unconscious processors. On this view, then, conscious mental states are identical to globally broadcasted contents. While conscious processing is slow, serial, and error-prone, unconscious processing is quick, massively parallel, and highly reliable. The advantages of conscious processing are that it supports creative problem solving in response to novel circumstances, goal-driven action planning, and complex learning.

While I agree with Carruthers that the phenomenal concept strategy is the best if not the only way to reconcile realism with physicalism,¹² I argue ahead that scientists should reject inferential independence. For Carruthers, however, inferential independence is the very thing that a successful scientific theory of consciousness must explain. In his view, it is in virtue of the fact that the global workspace theory explains inferential independence that it counts as an explanation of consciousness at all. In response to the worry that the global workspace cannot reductively explain consciousness because of the explanatory gap, Carruthers argues as follows:

The key reason for thinking that phenomenal consciousness doesn't just co-occur with global broadcasting, but that it is globally broadcast nonconceptual content, is that the resulting theory can then be extended to explain, in a scientifically acceptable way, why people find phenomenal consciousness so puzzling—indeed, it can explain why they should so frequently come to believe

¹ Block (2002), Chalmers (1996), Papineau (2002, 2003).

² Hammeroff and Penrose (1996, 2003), Koch (2004, 2012), Penrose (1989, 1994), Tononi (2008, 2012), and Tononi and Koch (2015) are sometimes taken to endorse panpsychism, a form of property dualism.

³ Some representative examples are Baars (1988, 1997), Damasio (1999, 2010), Dehaene (2014), Edelman (1992), Edelman and Tononi (2000), Feinberg and Mallatt (2016, 2018, 2019), Graziano (2013), Lamme (2006, 2010, 2018), McFadden (2002a, 2002b), Metzinger (2003, 2009), Prinz (2012), and Shanahan (2006).

⁴ Loar (1990, 1997), McLaughlin (2001), and Papineau (2002), Tye (2003).

⁵ Loar (1997) states explicitly that “Phenomenal concepts are conceptually irreducible in this sense: they neither a priori imply, nor are implied by, physical-functional concepts.” Papineau (2002, pp. 97–102 & pp. 179–180) argues that phenomenal concepts and physical/function concepts lack inferential connections to one another. Carruthers & Veillet (2007, p. 213) insists that phenomenal concepts “are *conceptually isolated* from any other concepts that we possess, lacking any a-priori connections with non-phenomenal concepts of any type (and in particular, lacking such connections with any physical, functional, or intentional concepts).” Block (1990, 1995) offers several arguments for the conclusion that the qualitative aspects of experience, or P-consciousness, lack conceptual connections to functional structure.

⁶ Loar (1990, 1997).

⁷ Papineau (2011) offers a similar characterization of phenomenal concepts.

⁸ Block (2002), Papineau (2002).

⁹ Carruthers (2018, p. 51).

¹⁰ Carruthers (2011, 2015, 2019).

¹¹ Baars (1988, 1997), Dehaene (2014), Edelman et al. (2011), Shanahan (2006).

¹² The best-known criticism of the phenomenal concept strategy is, perhaps, Chalmers' (2007) so-called “master argument.” Chalmers' argument, however, depends on the controversial assumption that successful scientific explanation must provide a set of truths that *a priori* entail the facts they reduce. In agreement with Block and Stalnaker (1999), I reject the view that identity statements in science are deducible *a priori* from reducing theories. Insofar as identity statements appear in science, they are defended on *a posteriori* abductive grounds.

that consciousness cannot be scientifically explained. This is because one can form purely indexical, acquaintance-based, concepts for the experiences in question that lack any intrinsic connections with physical, functional, or representational concepts.¹³

In other words, the global workspace theory counts a reductive theory of consciousness, not only because the contents of the global workspace perfectly correlate with the contents of phenomenal states, but also because it explains inferential independence. It explains the illusion that phenomenal properties are something other than globally broadcasted contents.

2. Carruthers' argument

Carruthers argues that the global workspace theory implies there are no facts of the matter about the distribution of animal consciousness.¹⁴ The crux of his argument is that there is a problematic mismatch between the global workspace and consciousness that leads to problems when the theory is extended to animals. Because, in Carruthers' view, the global workspace must be characterized relative to human cognitive capacities, some of which may be unique to humans, few if any animals fully instantiate its architecture. Instead, they only instantiate the global workspace to a degree. A natural way to apply the global workspace to animals, then, is to say that animals are only conscious to the extent that their minds instantiate the cognitive architecture described by the global workspace theory. The trouble with this suggestion, Carruthers argues, is that consciousness does not come in degrees. Although the contents of consciousness vary in their clarity and intensity, consciousness itself is binary. Either one is conscious, or one is not. Going forward, I will assume, like Carruthers, that the global workspace theory is true and consciousness does not come in degrees.

Carruthers argues that there are three ways to explain the mismatch between the global workspace and consciousness, and only one is viable. The three options are as follows:

- (i) Global broadcasting is either categorically present or categorically absent, just like consciousness.
- (ii) There is a vague boundary between animal minds that instantiate the global workspace in a way that is sufficient for consciousness and those do not.
- (iii) There are no facts of the matter about consciousness.

The first option seems the most straightforward solution, and the one that defenders of the phenomenal concept strategy would likely prefer. To categorically distinguish the animals that instantiate the global workspace from those that do not, the global workspace must be individuated by certain functions it performs. Carruthers packs a lot into his rejection of this idea, but his basic argument is simple: individuating the global workspace in terms of a narrower set of capacities could play no role in reductively explaining consciousness. Instead, "what does the explanatory work is the fact that the contents of the global workspace are made available to systems for reflective higher-order awareness and discursive thought. For it is this that gives rise to the 'hard problem' of consciousness, the conceivability of zombies, and so on." In other words, Carruthers falls back on his assumption that reductively explaining consciousness is just a matter of explaining why introspection seems to reveal that phenomenal concepts and physical/functional concepts are inferentially independent.

To reject the second option, Carruthers points out that there are different ways of weighting dimensions of similarity to the global workspace.¹⁵ Different weightings result in different verdicts about where the line between conscious animals and non-conscious animals must be drawn. Is it more important the contents of the global workspace are distributed to the systems responsible for higher-order thought, or is it more important that they are made available to the systems responsible for reasoning and decision making? Or does it suffice that they are made available to short-term memory? How these questions are answered will dramatically impact the distribution of animals that are deemed conscious. But the problem, according to Carruthers, is that there are no facts of the matter about which of these systems are more important for consciousness. Why are there no facts of the matter? Carruthers does not explicitly say, but the underlying reason is clear enough: phenomenal concepts and physical/functional concepts are inferentially independent of one another. He acknowledges that there may be some animals whose minds are so similar to human minds that they are counted as conscious on "any reasonable way of weighting the different components and dimensions."¹⁶ In that case, he is happy to allow that scientists may be justified in judging that these animals are conscious. But he insists that the number of such animals is almost certainly quite small and may not extend beyond humans.

With the first two options ruled out, Carruthers argues that the best explanation of the mismatch is that there are no facts of the matter about animal consciousness. The global workspace theory of consciousness successfully explains human consciousness, but it can render no verdict about animal consciousness. This is not because animal consciousness is unknowable, but rather because there are no facts of the matter to be known. All that can be known is that different animal minds exhibit varying degrees of similarity to the global workspace.

As stated, the argument only applies to the global workspace theory of consciousness. But supposing that the underlying argumentative strategy is sound, its full scope is much more expansive. Many scientific theories of consciousness are functionalist theories.¹⁷ That is, many of them attempt to explain consciousness in terms of certain cognitive capacities and their functional relations

¹³ Carruthers (2018, pp. 50-51).

¹⁴ Carruthers (2018, 2019).

¹⁵ Carruthers (2018, pp. 54-55).

¹⁶ Carruthers (2018, p. 55).

to other cognitive capacities, some of which are unique to humans or at most a small subset of animals. If that is right, then Carruthers' argument generalizes to many other functionalist accounts of consciousness. Consciousness is binary, but functionally specified cognitive systems come in degrees. Since any attempt to explain this mismatch would fail to explain why phenomenal concepts are inferentially independent, it appears the most reasonable conclusion is that there are no facts of the matter about whether they are conscious.

Carruthers' interpretation of the global workspace theory is far from orthodox. Proponents generally believe that consciousness extends to other animals and that their theory can explain why. This reflects a broader consensus among scientists that consciousness at least extends to other mammals.¹⁸ Baars is explicit on the matter¹⁹ and has co-authored numerous articles on animal consciousness and the evolutionary origins of consciousness.²⁰ Stanislas Dehaene, another prominent defender of the global workspace theory, relies heavily on animal studies to defend the global workspace theory and explicitly states that we share the global workspace in common with many other animals.²¹ While it is true that they often focus on the roles that the global workspace plays in informing uniquely human consumer systems, there is little reason to think that, in their view, the global workspace must be so characterized. Given their view on animal consciousness, the more charitable interpretation is that they believe the global workspace can be characterized in terms of a core set of functions, and a core set of consumer systems, that are shared in common with many other animals. Carruthers disagrees with them, of course, but it is conspicuous that he does not address the fact that his argument disagrees so radically with the way the theory is typically characterized by its major proponents.

If Carruthers is right that the global workspace must be singled out by cognitive capacities that are unique to humans, much of the evidence that is standardly given in support of the theory is undermined. This not only poses a problem for the way that the theory is typically defended, but also for Carruthers' idea that the main reason to accept the theory is that it can account for the explanatory gap. If the contents of the global workspace do not even correlate with the contents of phenomenal awareness, there would be little reason to think that the global workspace could explain why phenomenal concepts are inferentially independent.

Accordingly, scientists may be more inclined to reject his interpretation of the argument than to agree with his conclusion. To avoid this outcome, Carruthers must show that his interpretation is the best way, if not the only way, to make sense of the theory. This is a high bar.

3. Science requires there are facts about the distribution of consciousness

Supposing the global workspace theory is true, Carruthers' argument implies that there are no facts of the matter about animal consciousness. But I contend that the argument proves too much, for it also implies that there are no facts of the matter about the distribution of consciousness among humans. Supposing that Carruthers is right that global broadcasting must be characterized by cognitive capacities that are unique to humans, the question becomes, which humans, and which cognitive capacities? Young children and the cognitively disabled do not possess the same cognitive capacities as normal, healthy adults. Yet some young children and some cognitively disabled are conscious in the same way and to the same extent that healthy adult humans are conscious. This is not to beg any important questions or oversimplify the issue: it is an interesting and difficult question which cognitive capacities are required for conscious experience and which are not. Nor is it to criticize Carruthers for thinking that some humans lack consciousness, a position he previously defended.²² The problem is that, if scientists adopt Carruthers' interpretation of the global workspace, then they must conclude that there are no facts of the matter about whether or not children and the cognitively disabled are conscious because they do not possess the full catalog of cognitive capacities that normal human adults do. But surely an adequate theory of consciousness ought to position scientists to say at what point humans acquire consciousness and at what point they lose it. Dehaene thinks that a critical test of a theory of consciousness is whether or not it can help scientists to correctly distinguish persistent vegetative state from minimal awareness and total locked-in syndrome. Carruthers' interpretation of the global workspace theory is unable to provide such an answer.

So far, none of this is likely to persuade Carruthers, although it makes the tension between his interpretation of the global workspace and the way the theory is conceived by its major proponents all the more apparent. The more fundamental problem with Carruthers' view is that it implies there are no facts of the matter about human consciousness, period. Humans differ in their cognitive, perceptual, and motor capacities, both individually and across cultures.²³ Presumably, these differences reflect cognitive differences. If so, global broadcasting for me may be tied to different cognitive capacities than global broadcasting for you. How can I be sure, then, that you are conscious in the same way that I am, or indeed conscious at all? Employing the same sort of reasoning that Carruthers uses to rule out facts of the matter about animal consciousness, it follows that there are no facts of the matter about human consciousness either. Just as there is no determinate way to rank degrees of similarity between cognitive architecture that supports consciousness in humans and similar cognitive architecture in animals, so there is no determinate way to rank degrees of similarity

¹⁷ Arguably nearly all theories in cognitive science are some form or other of functionalist theory (Hardcastle 1996).

¹⁸ Most scientific leaders believe that many non-human animals are conscious. "The Cambridge Declaration on Consciousness," signed by many prominent scientists, reports that consciousness at least extends to mammals and birds, likely extends to all vertebrates, and possibly extends to arthropods and mollusks.

¹⁹ Baars (1997).

²⁰ See, for example, Boly et al. (2013), Seth et al. (2005).

²¹ Dehaene (2014), Dehaene & Naccache (2001).

²² Carruthers (2000).

²³ Gruszka et al. (2010), Henrich et al. (2010).

between the cognitive architecture that supports consciousness in me and similar cognitive architecture in you. I know that I am conscious in my own case, and I know that the cognitive architecture that supports consciousness in my own case is fully explanatory of consciousness in me, but I must conclude that there are no facts of the matter about whether people whose cognitive architectures differ from mine are truly conscious.

Carruthers would likely respond to this criticism by insisting that scientists can rely on first-person testimony to sort out which humans are conscious and what cognitive capacities they all share in common. But for this answer to work, it must also tell scientists who is *not* conscious. Carruthers cannot say that there are some humans for which there is no fact of the matter about whether or not they are conscious since the explanatory target of the global workspace theory gets decided by which humans are conscious and, more importantly, which humans are not. Perhaps Carruthers is right that it is possible to give a satisfactory theory of consciousness that only applies to humans. But it is not possible to give a satisfactory theory that only applies to *some* humans.

Ultimately, then, scientists must isolate a privileged set of capacities to individuate the global workspace that have definite implications for the distribution of consciousness, if not among animals in general, then at least among humans. If so, verbal report is no longer the only source of evidence that needs to be considered. We must also consider individuals who cannot report on their experiences: e.g., aphasics, prelingual children, children with prelingual deafness, minimally conscious patients, patients with total locked-in syndrome, and so on. Scientists have recently devoted considerable attention to these individuals²⁴ by taking advantage of experimental techniques like binocular rivalry²⁵ and masking,²⁶ as well as technologies like transcranial magnetic stimulation²⁷ that allow scientists to study the impact of brain injury on conscious experience.²⁸ Dehaene himself is at the forefront of some of this research. Not only does he use it to support the global workspace theory of consciousness, but he has also proposed certain neural markers for distinguishing minimally conscious patients from vegetative patients.²⁹

The claim that verbal report offers the only direct way to detect consciousness and learn about its content is highly contentious. Scientists have developed established methods for training animals to report non-verbally on their subjective experiences to test for blindsight in monkeys.³⁰ Although Carruthers criticizes the evidence for metacognition in animals,³¹ the methods in question do not rely on purported meta-cognitive capacities. Instead, they rely on first-order capacities to distinguish and react to different visual stimuli. Scientists have yet to isolate minimal neural correlates for consciousness,³² but there are several neural markers for consciousness that are regarded as reliable in both humans and other mammals, and scientists are beginning to arrive at similar markers for consciousness for non-mammalian animals.³³

Once it is acknowledged that scientists must go beyond verbal reports to determine the distribution of consciousness among humans, there is little reason to resist the idea that they should do the same for animals. Some animals, after all, are more clearly conscious than many human patients in a minimally conscious state. Yes, scientists may have to rely on fallible and incomplete information, but for reasons just discussed, determining the distribution of consciousness among humans faces similar uncertainty. Just as scientists can rely on developmental, neurophysiological, and behavioral similarities between verbal and nonverbal humans, so they can rely on such similarities between humans and animals. In addition, there are significant neurophysiological similarities between human brains and the brains of other mammalian animals that strongly suggest evolutionary continuity between animal and human cognition. It would be surprising if something as fundamental to human cognition as consciousness turned out to be a uniquely human trait. It would be analogous to discovering that the human immune system functions in a fundamentally different way than the immune systems of other mammals to which we are closely related on the phylogenetic tree.

Presently there is no consensus about which animals are conscious and which unconscious. But nearly all scientific theories of consciousness have implications for the distribution of consciousness. The point here is not that scientific theories of consciousness must establish the distribution of consciousness once and for all, but rather that they cannot avoid making commitments that have implications for the distribution of consciousness. Carruthers cannot escape the issue either. But since the conclusion of Carruthers' argument is that there are no facts of the matter about the distribution of consciousness, there must be something defective about the argument. Where does the argument go wrong? My view is that the argument goes wrong by presupposing the phenomenal concept strategy—a strategy that, arguably, scientists have good reason to be wary of. After all, many of its most prominent defenders argue that the phenomenal concept strategy may rule out the possibility of a legitimate science of consciousness.³⁴

4. Rejecting conventional versions of the phenomenal concept strategy

It should now be clear that the main pillar of support for Carruthers' argument is the phenomenal concept strategy. More

²⁴ Demertzi et al. (2019), Luppi et al. (2019), Owen et al. (2006).

²⁵ Tong et al. (1998).

²⁶ Dehaene et al. (1998).

²⁷ Baars et al. (2013), Jolij & Lamme (2005).

²⁸ Clark and Squire (1998).

²⁹ Dehaene (2014), Bekinschtein et al. (2009).

³⁰ Cowey and Stoerig (1995), Stoerig & Cowey (1995), Leopold and Logothetis (1996), Logothetis and Sheinberg (1996), Leopold and Logothetis (1999).

³¹ Carruthers (2008).

³² Koch et al. (2016).

³³ Boly et al. (2013), Bronfman et al. (2016), Edelman (2003), Feinberg and Mallatt (2019), Leopold et al. (2003), Seth et al. (2005).

³⁴ Block (2002), Papineau (2002).

narrowly, it is the claim that the global workspace theory must be paired with the phenomenal concept strategy to avoid the problems raised by the explanatory gap. Only when the global workspace theory is understood from the perspective of the phenomenal concept strategy, Carruthers argues, is it possible to explain why conscious mental states not only co-occur with globally distributed contents but are also identical to them. Similar points would apply to other functional theories of consciousness. A functionalist theory that could not explain why phenomenal concepts are inferentially independent of physical/functional concepts could not explain why there seems to be an explanatory gap. So, again, the argument generalizes.

It would seem, then, that scientists who think there are facts of the matter about animal consciousness should reject the phenomenal concept strategy. But things are not so simple. The phenomenal concept strategy remains the most promising strategy to make physicalism consistent with the explanatory gap. It remains, in my view, the only explanation for why, when we introspect on features of conscious states, we struggle to accept the possibility that they might be identical to physical/functional states. Rather than reject the phenomenal concept strategy wholesale, my view is that scientists should embrace the features of the phenomenal concept strategy that make realism compatible with physicalism while rejecting features that imply there are no facts of the matter about animal consciousness.

Part of what is at issue is the source of the explanatory gap. According to the conventional view, the source of the gap is that physical/functional facts do not entail any non-trivial phenomenal facts. But I agree with Papineau (2011) that the real source of the gap is simply the psychological disposition to reject identity statements between phenomenal states and physical/functional states. The explanation for this gap is a psychological question that, in all likelihood, will eventually be answered by science. Understood this way, one can endorse the phenomenal concept strategy as a response to the explanatory gap while denying that it serves as a reductive theory of consciousness. One can think that no physical/functional *reduction* of consciousness is possible while denying that phenomenal states are illusory and insisting that an adequate theory should describe the physical/functional processes to which consciousness is identical. To do this, scientists can side with Block and Stalnaker (1999), who argue that identity statements in science are usually if not always supported by abductive, *a posteriori* arguments. If so, a reductive theory of consciousness does not need to analytically entail the existence of phenomenal states and their associated features—in particular, such a theory does not need to explain inferential independence.

Since the explanatory gap seems to show that no physical/functional process can offer a satisfying explanation of consciousness, scientists support theories of consciousness with evidence of perfect correlation. Relying on perfect correlation, however, is problematic on the assumption that phenomenal concepts and physical/functional concepts are inferentially independent. Block, for example, argues that inferential independence implies there is no way to verify whether or not non-humans (including aliens and AI) are conscious.³⁵ In a similar vein, Papineau argues that inferential independence implies there is no way for theories of consciousness to select between minimal neural correlates and the higher-order functional states they instantiate.³⁶

Standardly, inferential independence is used to explain away the explanatory gap. For the most part, however, the arguments for the explanatory gap only run in one direction: they pose difficulties for inferences from physical/functional facts to phenomenal facts, but not vice versa. Of course, as Colin McGinn points out, there does not seem to be anything in conscious experience that resembles the physical parts that make up the brain.³⁷ He takes this to imply that there is no way to understand conscious mental phenomena via a physical theory since physical theories always take the form of analyses that explain the features of physical objects in terms of the combinatorial structure of their component elemental parts. While we also do not directly experience atoms, electrons, quarks, and so on, McGinn insists that introspection on consciousness is unlikely to support similarly complex theories about the combinatorial structure of conscious experience. So, it seems, the physical basis of consciousness must remain a mystery.

The solution to these problems, in my view, is to reject total inferential independence. If scientists knew something about the functional structure of consciousness prior to empirical investigation, they could use that knowledge to isolate the minimal neural correlates of consciousness and evaluate among competing theories. While there are no non-trivial inferential connections from physical/functions truths to phenomenal truths, and while there are no inferential connections from phenomenal truths to strictly physical truths, there are inferential connections from phenomenal truths to functional truths. If that is right, then uncovering the physical/functional basis of consciousness may not be much different from finding the physical/functional basis of any other natural phenomenon: the goal is to offer a theoretical explanation of certain observable causal patterns in terms of the lawlike behavior of unobservables. The epistemic problems raised by Block and Papineau would largely disappear.

A further reason to prefer my version of the strategy over conventional versions is that it avoids the tendency for the phenomenal concept strategy to collapse into property dualism. Setting aside deflationary versions of the phenomenal concept strategy,³⁸ there is reason to doubt that we can make sense out of the idea that the essential nature of phenomenal properties is opaque to us even though they are directly grasped by immediate acquaintance. As Demircioglu argues, “A non-deflationist physicalism is a dead-end, whose failure brings about an important lesson: the recognition of a category of concepts of phenomenal experiences whose cognitive contents bear no *a priori* links to the cognitive contents of physical concepts is fatal to physicalism.”

Alternatively, the version of the strategy that I am proposing points to a middle road between absolute inferential independence and deflationism. It suggests that phenomenal concepts are *partly* dependent on physical concepts. One can deny that phenomenal concepts are completely independent of physical concepts while recognizing that there are aspects of our phenomenal concepts that

³⁵ Block (2002).

³⁶ Papineau (2002).

³⁷ McGinn (1993, 1996, 1999).

³⁸ Levin (2007).

are not susceptible to *a priori* analysis. Importantly, then, my version of the strategy does not endorse functionalism about consciousness: one can believe that consciousness has necessary functional structure while remaining agnostic about the possibility that consciousness is nothing over and above functional structure. Nor does it endorse identity theory. The idea is that phenomenal states are identical to physical/functional states that act as their own modes of presentation but do not make their complete essential natures transparent. Although they reveal something about their functional structure, they do not reveal anything substantive about the neural processes that support them.

Some might balk at the suggestion that my view is consistent with the phenomenal concept strategy. If phenomenal concepts are partly dependent on physical/functional concepts, then it might appear that a reductive account of consciousness should be possible. If so, then there should be no explanatory gap between physical/functional concepts and phenomenal concepts. (To use Chalmers' well-known distinction, my approach would become a version of type-A physicalism rather than type-B physicalism.³⁹) But the reality is that my account rules out reduction, whether analytic or scientific. It agrees that there are no non-trivial inferential connections from physical/functional concepts to phenomenal concepts. Accordingly, no scientific reduction of phenomenal truths in terms of physical/functional truths is possible. There is no physical/functional theory of consciousness that entails the existence of phenomenal truths from the existence of human nervous systems.

The suggestion is that there are, nonetheless, inferential connections from phenomenal concepts to physical/functional concepts that can position scientists to isolate a core set of functions played by phenomenal concepts. These inferential connections are, however, partial: while they do not allow scientists to draw conclusions about the broad functional structure of any physical system that instantiates consciousness, they do not position scientists to infer detailed conclusions about its physical/functional nature. For that, science is required. But they can help to guide theory selection by offering a set of functional desiderata that any theory must satisfy. Most importantly, in the present case, they can avoid the implication that there are no facts of the matter about which capacities of the physical/functional system that realizes consciousness are more important, or more central, to the nature of consciousness than others.

To sum up, I have suggested that it is possible to offer a well-motivated theory of consciousness that does not rely on explaining away the explanatory gap. Perfect correlation can also do the job. For this strategy to be successful, however, scientists must reject inferential independence. In the next section, I offer some reason to think that inferential independence is false.

5. Consciousness offers a unified map that informs decision-making processes

My view is that consciousness has necessary functional structure that we can ascribe to conscious mental states on the basis of introspection. While certain aspects of experience taken in isolation may not seem to possess any functional structure (e.g., a particular shade of red), consciousness experience as a whole is rife with it. I argue that conscious experience offers a persistently updated, unified map of the organism in its environment that is made available in real time to consumer systems responsible for decision making. To be clear, this characterization is neither meant to serve as a definition of consciousness nor a theoretical account of consciousness. It offers necessary conditions, but not sufficient conditions for conscious experience. Nor is it meant to serve as a constraint on theory-construction, since most theories are already consistent with it. But if true, it rules out bidirectional inferential independence and it points to a core set of functions that any adequate theory of consciousness must explain.

The main piece of introspective support for this functional characterization of consciousness is the so-called transparency of experience. When one tries to introspect on the subjective features of one's experience, one's attention is inexorably drawn to features and states of one's body or external objects in the environment.⁴⁰ One seems to *see through* experience. Many philosophers argue that the best explanation for transparency is representationalism, the view that conscious mental states have intentional content.⁴¹ Conscious experience directs our attention away from itself and towards external objects and properties because it is about something. While the weaker forms of representationalism enjoy widespread support, there are some dissenters, and their arguments are not easily dismissed. Ned Block, for example, argues that there may be features of phenomenal states (i.e., qualia) that lack representational content altogether.⁴² For present purposes, I am willing to grant this point. I will insist, however, that *most* phenomenal states have necessary representational content: that some phenomenal states could not be the phenomenal states that they are in the absence of certain representational content. We might call this view minimal representationalism. Minimal representationalism is endorsed (in some cases implicitly) by the vast majority of contemporary theorists working in consciousness studies, including Block.

By itself, minimal representationalism does not imply anything about the functional structure of consciousness. To be of use to scientists, it must be paired with certain views about the determination of representational content. Nearly all philosophical theories of content determination are committed to the idea that mental representations must play some or other functional roles to have the representational content that they do.⁴³ Mental representations do not have their contents as a matter of brute fact. At a minimum, all theories of content determination are committed to the idea that the contents of mental representations must be available to certain

³⁹ Chalmers (1996, 2007).

⁴⁰ Harman (1990), Tye (1995).

⁴¹ See Tye (2002) for a defense of strong representationalism. See Block (2002), Chalmers (1996, 2004), Loar (2003) for discussion about weak representationalism.

⁴² Block (1996, 2003, 2010).

⁴³ See, for example, Dretske (1995), Fodor (1987, 1990a, 1990b, 1994), and Millikan (1984).

goal-driven consumer systems. If that is right, then we can conclude that the content of conscious mental states must be available to certain consumer systems. This has significant functional implications that, when combined with unity and valence of experience, place strong functional constraints on scientific theories of consciousness.

The unity of consciousness has many different sources, but most of them have functional implications.⁴⁴ The main sources are temporal unity, objecting binding, and integration.

Temporal unity: Moments of conscious experience are structured sequentially by relationships of simultaneity and succession. How one experiences the present time is always informed by the recent past and expectations about the immediate future. In this way, the information carried by conscious experience is spread out across time.

Object binding: The representational contents of one's conscious experiences are essentially integrated. One type of representational unity is aimed at representing particular objects. Presently, a half-eaten apple is resting on my desk. Somehow, my experience of its redness is bound together with my experience of its shape, its smell, its taste, and so on, so that they all belong to one and the same object. Explaining how the brain achieves this kind of representational unity is sometimes referred to as the binding problem.⁴⁵ Although scientists disagree about how to resolve the problem, it offers a functional desideratum on scientific theories of consciousness.

Integration: Under ordinary circumstances, one is consciously aware of many things all at once: different objects, different properties, different states of one's body, and so on. All of these contents come together to form a present-centered scene of one's environment as it relates to one's body. While it is possible to distinguish these different elements of experience, their phenomenal character and representational contents are essentially tied to one another. My experience of the half-eaten red apple in front of me, for example, is informationally integrated with my conscious experience of the position of my body and how it appears in relation to other objects on my desk. Although my conscious experience of the apple is distinct from my conscious experience of my body, the content of either experience is informationally connected to the content of the other. I could not have the same phenomenal experience of either one if they were not representationally unified.

Lastly, conscious experience is often colored by positive and negative valence, internal drives that motivate us to seek out nourishment, avoid tissue damage, maintain chemical homeostasis, and so on. Under normal circumstances, conscious experience is teeming with various affective qualities that motivate the experienter to perform (or avoid) certain actions. Conscious hunger disposes one to seek out and consume food, conscious pain disposes one to withdraw from the perceived source of the pain and attend to any tissue damage, conscious mirth disposes one to laugh, and so on. Arguably, these phenomenal qualities could not be the qualities that they are absent from their various motivational affects. As Chalmers notes, the dispositional connections between affective states and their corresponding behaviors are so tightly connected to one another that our ordinary terms for referring to affective states tend to conflate them with the behaviors they dispose.⁴⁶ While Chalmers is right that we can make the conceptual distinction between affective states and their corresponding dispositions, it is no surprise that this conflation is so ubiquitous: many affective states are so tightly linked to their corresponding dispositions that we cannot make clear conceptual sense out of pulling them apart.

We can, of course, distinguish between the strictly phenomenal aspects of affective states and their accompanying dispositions, but that does not mean that we can conceive of them states in the total absence of their dispositions. Even Putnam's *super-Spartans* (1963), humans that have trained themselves to behave as if they never experience pain, still experience the urge to behave as if they are in pain and still require training to resist those urges. Super-Spartans represent a good counterexample to classical behaviorism, but they pose no great difficulty for the idea that pain necessarily motivates pain behavior. A more serious difficulty seems to be suggested by Strawson's *hypothetical weather watchers* (1994). Weather watchers are alien creatures that are completely immobile and altogether incapable of any form of behavior. Nevertheless, they consciously experience the local weather and they have various beliefs, thoughts, desires about the weather. Setting aside worries about the conceivability of such creatures, they do not threaten the idea that affective states are necessarily linked to certain dispositions. The affective states of weather watchers still have functional implications for their future mental states. They may feel elated when the weather turns out the way that they hoped, and they may feel despair when it does not. More to the point, their affective states still imply certain behavioral dispositions were the creatures in question to acquire capacities for behavior. The conceivability of weather watchers may force us to acknowledge that there are no *unconditional* links between affective states, motivation, and behavior. But they are consistent with the idea that affective states imply conditional dispositions.

Like the unity of experience, the valence of experience has implications for the functional structure of consciousness when paired with minimal representationalism.⁴⁷ At a minimum, it implies that enough information is passed along to the relevant consumer systems to motivate the relevant behaviors. This includes not only the positive or negative valence but also the represented contents of whatever that valence attaches to. Since, as I have already argued, much of representational content of experience is holistically integrated, this strongly suggests that the entire contents of one's experience, along with its myriad affective aspects, must be made available to the relevant decision-making processes.

Put together, the intentionality, unity, and valence of conscious experience tell scientists that the cognitive system that

⁴⁴ See Bayne (2010) for more on the unity of conscious experience.

⁴⁵ Treisman (1996).

⁴⁶ Chalmers (1996, pp. 17-18).

⁴⁷ Neuroscientist Jaak Panksepp (1998) has arguably made extensive use of these functional implications. He has worked extensively on the functional and motivational structure of affective consciousness and its implications for the affective systems of animal nervous systems.

implements conscious experience must take the form of a persistently updated, present-centered, integrated, internally consistent map of the animal in its environment that is made available to a variety of consumer systems, especially those responsible for decision-making. There are surely other sources of introspectable functional structure that I have not touched on. But let us stop here, as the point has been illustrated well enough: although some philosophers have seemed to deny it, consciousness implies functional structure. Even if the reader is unpersuaded by some of the details of this section, or even if she rejects, for example, that experience is necessarily valenced, so long as she's been persuaded of the general point that conscious experience has necessary functional structure, then she agrees that inferential independence should be rejected.

To wrap up this section, I should stress that all of the arguments given above describe necessary functional characteristics of *phenomenally* conscious states. This is important to emphasize because some philosophers, notably Block (1995), argue that scientists regularly conflate different notions of consciousness in a way that leads to fallacious conclusions.⁴⁸ But the transparency argument for representationalism specifically targets the phenomenal character of conscious experience. While the discussion on the unity of consciousness mostly targeted its representational unity, it is relevant to understand the functional structure of phenomenal character so long as minimal representationalism is true. And while it is possible to distinguish the phenomenal character of conscious affects from the behavior it motivates, it is essential to the phenomenal character of affective states that they motivate the behavior that they do.

6. Conclusion

We have arrived at an introspectable characterization of the functional structure of consciousness. I note again that this characterization should not be understood as *a theory* of consciousness since it does not attempt to provide sufficient conditions for conscious experience. Instead, it describes a set of necessary functions that any adequate theory of consciousness should account for. This offers an alternative way to defend theories of consciousness that does not rely on bridging the explanatory gap. Instead, the goal is to find the cognitive systems in animal brains capable of playing the requisite functional roles. In many ways, this makes theorizing about consciousness not so different from theorizing about anything else. Like much of physics, chemistry, and biology, scientists investigating the basic nature of consciousness rely on observable causal patterns to theorize about an underlying world of unobservable properties and laws.

Given what I have argued, either of the two alternatives considered by Carruthers' argument is preferable to the view that there are no facts of the matter about consciousness. Carruthers rejects the idea that scientists might individuate the global workspace in terms of its basic functions because doing so would contribute nothing to explaining inferential independence, which, in his view, is the only way to infer identity from correlation. But I have argued that scientists can infer identity from correlation by appealing to pre-theoretical introspective knowledge about the functional structure of consciousness. Similarly, Carruthers rejects the idea that there might be a vague boundary between the conscious animals and the unconscious animals on the grounds that there are different ways of weighting the similarity space but no facts of the matter that could decide which ranking should be preferred. But if consciousness has necessary functional structure, then there are facts of the matter about which consumer systems and which features of the global workspace are more critical than others.

Author statement

Dylan Black is the sole author the attached manuscript.

Acknowledgments

I would like to express my appreciation to Colin Allen, Kirk Ludwig, Gary Ebbs, and Kevin Mills for their feedback and constructive criticism.

References

- Baars, B. J. (1988). *A cognitive theory of consciousness*. Cambridge: Cambridge University Press.
- Baars, B. (1997). *In the theater of consciousness*. New York: Oxford University Press.
- Baars, B. J., Franklin, S., & Ramsøy, T. Z. (2013). Global workspace dynamics: Cortical "binding and propagation" enables conscious contents. *Frontiers in Psychology*, 4, 200.
- Bayne, T. (2010). *The unity of consciousness*. Oxford: Oxford University Press.
- Bekinschtein, A., Dehaene, S., Rohaut, B., Tadel, F., Cohen, L., & Naccache, L. (2009). Neural signature of the conscious processing of auditory regularities. *Tristan Proceedings of the National Academy of Sciences*, 106(5), 1672–1677.
- Block, N. 1990. "Inverted Earth," *Philosophical Perspectives*, 4, J. Tomberlin, ed., Atascadero, CA: Ridgeview Publishing Company.
- Block, N. (1995). On a confusion about a function of consciousness. *Brain and Behavioral Sciences*, 18(2).
- Block, N. (1996). Mental paint and mental latex. *Philosophical Issues*, 7, 19–49.
- Block, N. (2002). The harder problem of consciousness. *Journal of Philosophy*, 99(8), 391.
- Block, N. (2003). Mental paint. In M. Hann, & B. Ramberg (Eds.). *Reflections and replies: Essays on the philosophy of Tyler burge* (pp. 165–200). MIT Press.
- Block, N. (2010). Attention and mental paint. *Philosophy Issues*, 20, 23–63.
- Block, N., & Stalnaker, R. (1999). Conceptual analysis, dualism, and the explanatory gap. *Philosophical Review*, 108(1), 1–46.

⁴⁸ See Block (1995), for more discussion on this point.

- M. Boly A.K. Seth M. Wilke P. Ingmundson B. Baars S. Laureys ... N. Tsuchiya Consciousness in humans and non-human animals: Recent advances and future directions *Frontiers in Psychology* 4 2013 625.
- Bronfman, Z. Z., Ginsburg, S., & Jablonka, E. (2016). The Transition to Minimal Consciousness through the Evolution of Associative Learning. *Frontiers in Psychology*, 7, 1954.
- Carruthers, P. (2000). *Phenomenal consciousness: A naturalistic theory*. Cambridge University Press.
- Carruthers, P. (2008). Metacognition in animals: A skeptical look. *Mind & Language*, 23(1), 58–89.
- Carruthers, P. (2011). *The opacity of mind*. New York: Oxford University Press.
- Carruthers, P. (2015). *The centered mind: What the science of working memory shows us about the nature of human thought*. New York: Oxford University Press.
- Carruthers, P. (2018). Comparative psychology without consciousness. *Consciousness and Cognition*, 63, 47–60.
- Carruthers, P. (2019). *Human and animal minds: The consciousness questions laid to rest*. Oxford University Press.
- Carruthers, P., & Veillet, B. (2007). The phenomenal concept strategy. *The Journal of Consciousness Studies*, 14(9–10), 212–236.
- Chalmers, D. (1996). *The conscious mind*. Oxford: Oxford University Press.
- Chalmers, D. (2004). The representational character of experience. In B. Leiter (Ed.). *The future for philosophy*. Oxford: Oxford University Press.
- Chalmers, D. (2007). Phenomenal concepts and the explanatory gap. In T. Alter, & D. Walter (Eds.). *Phenomenal concepts and phenomenal knowledge: New essays on consciousness and physicalism*. Oxford: Oxford University Press.
- Clark, R. E., & Squire, L. R. (1998). Classical conditioning, awareness, and brain systems. *Trends in Cognitive Science*, 6(12), 524–531.
- Cowey, A., & Stoerig, P. (1995). Blindsight in monkeys. *Nature*, 373, 247–249.
- Damasio, A. (1999). *The feeling of what happens: Body and emotion in the making of consciousness*. Harcourt Brace and Company.
- Damasio, A. (2010). *Self comes to mind: Constructing the conscious brain*. New York: Pantheon/Random House.
- Dehaene, S. (2014). *Consciousness and the brain*. New York: Viking Press.
- Dehaene, S., & Naccache, L. (2001). Towards a cognitive neuroscience of consciousness: Basic evidence and a workspace framework. *Cognition*, 79(1–2), 1–37.
- Dehaene, S., Naccache, L., Le Clec'H, G., Koehlin, E., Mueller, M., Dehaene-Lambertz, G., ... Le Bihan, D. (1998). Imaging unconscious semantic priming. *Nature*, 395, 597 ± 600.f.
- Demertzi, A., Tagliazucchi, E., Dehaene, S., Deco, G., Bartfeld, P., Raimondo, F., ... Voss, H. U. (2019). Human consciousness is supported by dynamic complex patterns of brain signal coordination. *Science Advances*, 5(2).
- Dretske, F. (1995). *Naturalizing the mind*. Cambridge, Mass.: The MIT Press.
- Edelman, G. M. 1992. Bright air, brilliant fire; on the matter of the mind. 1992. Basic Books.
- Edelman, G. M. 2003. Naturalizing consciousness: A theoretical framework. *Proceedings of the National Academy of Sciences*, 100(9), 5520 LP – 5524.
- Edelman, G., Gally, J., & Baars, B. (2011). Biology of consciousness. *Frontiers in Psychology*, 2, 4.
- Edelman, G. M., & Tononi, G. (2000). *A universe of consciousness: How matter becomes imagination*. Basic Books.
- Feinberg, T. E., & Mallatt, J. M. (2016). *The ancient origins of consciousness: How the brain created experience*. MIT Press.
- Feinberg, T. E., & Mallatt, J. M. (2018). *Consciousness demystified*. MIT Press.
- Feinberg, T. E., & Mallatt, J. M. (2019). Subjectivity “Demystified”: Neurobiology, Evolution, and the Explanatory Gap. *Frontiers in Psychology*, 10, 1686.
- Fodor, J. (1987). *Psychosemantics*. Cambridge, Mass.: The MIT Press.
- Fodor, J. (1990a). *A theory of content and other essays*. Cambridge, Mass.: The MIT Press.
- Fodor, J. (1990b). Psychosemantics or: Where do truth conditions come from? In W. G. Lycan (Ed.). *Mind and cognition: A reader* (pp. 312–337). Oxford: Blackwell Publishers.
- Fodor, J. (1994). *The elm and the expert*. Cambridge, Mass.: The MIT Press.
- Graziano, M. S. (2013). *Consciousness and the social brain*. Oxford: Oxford University Press.
- Gruzka, A., Matthews, G., & Szymura, B. (2010). Handbook of individual differences in cognition. *Attention, Memory and Executive Control* (pp. 87–107). Springer.
- Hammeroff, S., & Penrose, R. (1996). Orchestrated reduction of quantum coherence in brain microtubules: A model for consciousness. *Mathematics and Computers in Simulation*, 40, 453–480.
- Hammeroff, S., & Penrose, R. (2003). Conscious events as orchestrated space-time selections. *NeuroQuantology*, 1(1), 10–35.
- Hardcastle, V. G. (1996). *How to build a theory in cognitive science*. SUNY Press.
- Harman, G. (1990). The intrinsic quality of experience. *Philosophical Perspectives*, 4, 31–52.
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, 33(2–3), 61–83.
- Jolij, J., & Lamme, V. A. (2005). Repression of unconscious information by conscious processing: Evidence from affective blindsight induced by transcranial magnetic stimulation. *Proceedings of the National Academy of Sciences*, 102(30), 10747–10751.
- Koch, C. (2004). *The quest for consciousness: A neurobiological approach*. Roberts and Company.
- Koch, C. (2012). *Consciousness: Confessions of a romantic reductionist*. MIT Press.
- Koch, C., Massimini, M., Boly, M., & Tononi, G. (2016). “Neural correlates of consciousness: Progress and problems”: Erratum. *Nature Reviews Neuroscience*, 17(5), 307.
- Lamme, V. A. F. (2018). Challenges for theories of consciousness: seeing or knowing, the missing ingredient and how to deal with panpsychism. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 373(1755), 20170344.
- Lamme, V. A. F. (2006). Towards a true neural stance on consciousness. *Trends in Cognitive Sciences*, 10(11), 494–501.
- Lamme, V. A. F. (2010). How neuroscience will change our view on consciousness. *Cognitive Neuroscience*, 1(3), 204–220.
- Leopold, D., & Logothetis, N. K. (1996). Activity changes in early visual cortex reflect monkeys’ percepts during binocular rivalry. *Nature*, 379(6565), 549–553.
- Leopold, D., & Logothetis, N. K. (1999). Multistable phenomena: Changing views in perception. *Trends in Cognitive Sciences*, 3(7), 254–264.
- Leopold, D., Maier, A., & Logothetis, N. K. (2003). Measuring subjective visual perception in the nonhuman primate. *Journal of Consciousness Studies*, 10(9–10), 115–130.
- Levin, J. (2007). What is a phenomenal concept? In T. Alter, & S. Walter (Eds.). *Phenomenal concepts and phenomenal knowledge* (pp. 87–111). New York: Oxford University Press.
- Loar, B. (1990). Phenomenal states. *Philosophical Perspectives*, 4, 81–108.
- Loar, B. (1997). Phenomenal states II. In Ned Block, Owen Flanagan, & Güven Güzeldere (Eds.). *The Nature of Consciousness: Philosophical Debates*. MIT Press.
- Loar, B. (2003). Phenomenal intentionality as the basis of mental content. In M. Hahn, & B. Ramberg (Eds.). *Reflections and replies: Essays on the philosophy of Tyler Burge*. Cambridge: The MIT Press.
- Luppi, A. I., Craig, M. M., Pappas, I., Finoia, P., Williams, G. B., Allanson, J., ... Stamatakis, E. A. (2019). Consciousness-specific dynamic interactions of brain integration and functional diversity.
- Logothetis, N. K., & Sheinberg, D. L. (1996). Visual object recognition. *Annual Review of Neuroscience*, 19(1), 577–621.
- McLaughlin, B. (2001). In defense of new wave materialism. In C. Gillett, & B. Loewer (Eds.). *Physicalism and its discontents*. New York: Cambridge University Press.
- McFadden, J. (2002a). Synchronous Firing and its influence on the brain’s electromagnetic field. *Journal of Consciousness Studies*, 9(4), 23–50.
- McFadden, J. (2002b). The conscious electromagnetic information (cemi) field theory: The hard problem made easy? *Journal of Consciousness Studies*, 9(8), 45–60.
- McGinn, C. (1993). *Problems in philosophy: The limits of inquiry*. Massachusetts: Blackwell Publishers.
- McGinn, C. (1996). *The character of mind: An introduction to the philosophy of mind*. New York: Oxford University Press.
- McGinn, C. (1999). *The mysterious flame: Conscious minds in a material world*. New York: Basic Books.
- Metzinger, T. (2003). *Being no one: The self-model theory of subjectivity*. Cambridge: MIT Press.
- Metzinger, T. (2009). *The ego tunnel: The science of the mind and the myth of the self*. New York: Basic Books.
- Millikan, R. (1984). *Language, Thought and Other Biological Categories*. Cambridge, Mass.: The MIT Press.
- Owen, A. M., Coleman, M. R., Boly, M., Davis, M. H., Laureys, S., & Pickard, J. D. (2006). Detecting awareness in the vegetative state. *Science*, 313(5792), 1402.
- Panksepp, J. 1998. Affective neuroscience: The foundations of human and animal emotions. *Affective Neuroscience: The Foundations of Human and Animal Emotions*,

- pp. xii, 466–xii, 466. New York, NY, US: Oxford University Press.
- Papineau, D. (2002). *Thinking about consciousness*. Oxford: Clarendon Press.
- Papineau (2011). What exactly is the explanatory gap? *Philosophia*, 35, 5–19.
- Papineau, D. (2003). Could there be a science of consciousness? *Philosophical Issues*, 13(1), 205–220.
- Penrose, R. (1989). *The emperor's new mind: Computers, minds and the laws of physics*. Oxford: Oxford University Press.
- Penrose, R. (1994). *Shadows of the mind*. Oxford: Oxford University Press.
- Putnam, Hilary. Brains and Behavior. Mind, Language, and Reality: Philosophical Papers, Vol. 2: 325–341. Cambridge: Cambridge University Press. First appeared in 1963.
- Prinz, J. (2012). *The conscious brain: How attention engenders experience*. Oxford University Press.
- Seth, A. K., Baars, B. J., & Edelman, D. B. (2005). Criteria for consciousness in humans and other mammals. *Consciousness and Cognition*, 14, 119–139.
- Shanahan, M. (2006). A cognitive architecture that combines internal simulation with a global workspace. *Consciousness and Cognition: An International Journal*, 15(2), 433–449.
- Stoerig, P., & Cowey, A. (1995). Visual perception and phenomenal consciousness. *Behavioural brain research*, 71(1–2), 147–156.
- Strawson, G. (1994). *Mental reality*. Cambridge: MIT Press.
- Tong, F., Nakayama, K., Vaughan, J. T., & Kanwisher, N. (1998). Binocular rivalry and visual awareness in human extrastriate cortex. *Neuron*, 21, 753–759.
- Tononi, G. (2008). Consciousness as integrated information: A provisional manifesto. *The Biological Bulletin*, 215(3), 216–242.
- Tononi, G. (2012). Integrated information theory of consciousness: An updated account. *Archives Italiennes de Biologie*, 150(2–3), 56–90.
- Tononi, G., & Koch, C. (2015). Consciousness: here, there and everywhere? *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370(1668).
- Treisman, A. (1996). The binding problem. *Current Opinion in Neurobiology*, 6, 171–178.
- Tye, M. (1995). *Ten problems of consciousness*. Cambridge, MA: MIT Press.
- Tye, M. (2002). Representationalism and the transparency of experience. *Noûs*, 36, 137–151.
- Tye, M. (2003). A theory of phenomenal concepts. In Anthony O'Hear (Ed.). *Minds and Persons*. Cambridge University Press.