Special Issue: Consciousness science and its theories

Functions of consciousness: conceptual clarification

Takuya Niikawa[1,†], Katsunori Miyahara[2,*,‡], Hiro Taiyo Hamada[3,4,5,§], and Satoshi Nishida[6,7,‖]

[1]Graduate School of Humanities, Kobe University, 1-1 Rokkodai-machi, Nada-ku Kobe, Hyogo 657-8501, Japan; [2]Center for Human Nature, Artificial Intelligence, and Neuroscience (CHAIN), Hokkaido University, Kita 12 Nishi 7, Kita-ku, Sapporo, 060-0812, Japan; [3]Neurotechnology R&D Unit, Araya Inc., 1-12-32 Akasaka, Minato-ku, Tokyo 107-6024, Japan; [4]Department of Functional Brain Imaging, National Institute of Radiological Sciences, National Institutes for Quantum Science and Technology, 4-9-1 Anagawa, Inage-ku, Chiba 263-0024, Japan; [5]Neural Computation Unit, Okinawa Institute of Science and Technology, Graduate University, 1919-1 Tancha, Onna-son, Okinawa 904-0495, Japan; [6]Center for Information and Neural Networks (CiNet), Advanced ICT Research Institute, National Institute of Information and Communications Technology (NICT), 1-4 Yamadaoka, Suita, Osaka 565-0871, Japan; [7]Graduate School of Frontier Biosciences, Osaka University, 1-3 Yamadaoka, Suita, Osaka 565-0871, Japan

[†]Takuya Niikawa, http://orcid.org/0000-0002-4078-4917
[‡]Katsunori Miyahara, http://orcid.org/0000-0001-7747-6654
[§]Hiro Taiyo Hamada, http://orcid.org/0000-0002-7919-9125
[‖]Satoshi Nishida, http://orcid.org/0000-0003-0555-518X
The primary authorship of this paper is shared between TN and KM.
*Correspondence address. Center for Human Nature, Artificial Intelligence, and Neuroscience (CHAIN), Hokkaido University, Kita 12 Nishi 7, Kita-ku, Sapporo, Hokkaido 060-0812, Japan. Tel: +81 11 706 4897; Fax: +81-11-706-4049; E-mail: kmiyahara@chain.hokudai.ac.jp

Abstract

There are many theories of the functions of consciousness. How these theories relate to each other, how we should assess them, and whether any integration of them is possible are all issues that remain unclear. To contribute to a solution, this paper offers a conceptual framework to clarify the theories of the functions of consciousness. This framework consists of three dimensions: (i) target, (ii) explanatory order, and (iii) necessity/sufficiency. The first dimension, target, clarifies each theory in terms of the kind of consciousness it targets. The second dimension, explanatory order, clarifies each theory in terms of how it conceives of the explanatory relation between consciousness and function. The third dimension, necessity/sufficiency, clarifies each theory in terms of the necessity/sufficiency relation posited between consciousness and function. We demonstrate the usefulness of this framework by applying it to some existing scientific and philosophical theories of the functions of consciousness.

Keywords: function of consciousness; functional basis; functional contribution; meta-theoretical project

Introduction

The function of consciousness has been the focus of numerous recent scientific and philosophical theories of consciousness (Rosenthal 2008; Lau 2009; Cohen and Dennett 2011; Morsella and Andrew Poehlman 2013; Pierson and Trout 2017; Dennett 2018; Kanai *et al.* 2019; Birch *et al.* 2020; Wiese 2020; Black 2021). One reason for this is that it is common in both philosophy of mind and cognitive science to explore the place of the mind in the physical world by considering its functions. On these accounts, we can explain the physical basis of a conscious state by identifying its functional role in the cognitive economy and then specifying the physical mechanism that implements it (Dehaene 2014).[1]

Furthermore, assumptions about the functions of consciousness are often at work, either explicitly or implicitly, in discussions about the distribution and evolution of consciousness. When we consider which animal species have consciousness, our conclusions tend to be heavily influenced by considerations of whether they are equipped with certain cognitive functions (Ginsburg and Jablonka 2019). This suggests that specifying the functions of consciousness will be a crucial step toward determining how widely distributed consciousness is in the animal world and also understanding when and why it evolved.

To date, however, there is little consensus about the functions of consciousness (see next section for more discussion).

---

[1] This focus on function may not be necessary, or even the best way, to explore the relationship between the mind and its physical basis. For example, one might think it more promising to explore the isomorphism between the causal structure of the physical system and the phenomenological structure of conscious experience (Tononi and Koch 2015).
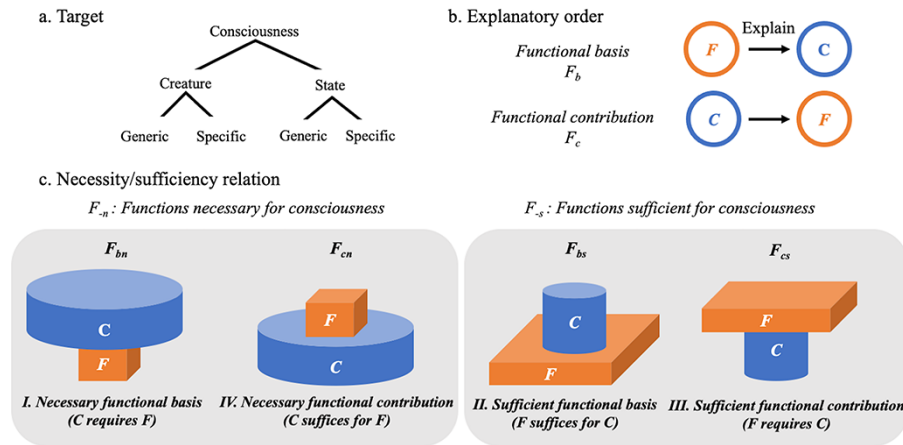
Figure 1. The three-dimensional framework consists of three dimensions: target, explanatory order, and necessity/sufficiency. (a) Target: four possible kinds of consciousness that can form the explanatory target of a theory of the functions of consciousness. (b) Explanatory order: two possible ways a theory can consider the explanatory relation between consciousness and its function. (c) Necessity/sufficiency: four possible ways in which a theory can envision the explanatory order and the necessity/sufficiency relation between consciousness and its function

Furthermore, the relationship between different theories of the functions of consciousness is often obscure. They do not always appear to pursue the same goal. Sometimes, they even seem to be based on different conceptions of functions and consciousness. Theories of consciousness itself are caught in a similar predicament. There is no consensus about what consciousness is, and the relationship between different theories of consciousness is far from clear. To resolve this troublesome situation, some researchers have developed approaches to systematically compare and assess theories of consciousness (Niikawa 2020; Doerig *et al.* 2021; Pin *et al.* 2021; Signorelli *et al.* 2021; Sattin *et al.* 2021). A similar meta-theoretical approach seems necessary to make progress in the study of the functions of consciousness.

The purpose of this paper is to contribute to this meta-theoretical project by developing a systematic framework for theories of the functions of consciousness (Fig. 1). This framework consists of three dimensions: (i) target, (ii) explanatory order, and (iii) necessity/sufficiency. The first dimension, target, locates theories in the framework based on the kind of consciousness that forms their explanatory target (Fig. 1a). For example, we can distinguish between theories that explore the functions of state consciousness and creature consciousness, the functions of perceptual experience and affective experience, and so on. The second dimension, explanatory order, locates theories based on their underlying conception of the explanatory relation between consciousness and functions (Fig. 1b). Some theories seek to explain consciousness in reference to its functions, while others seek to explain certain functions in reference to consciousness. Still others hold that consciousness is ontologically identical to a certain function, suggesting the absence of an explanatory order between them. The third dimension, necessity/sufficiency, locates theories in terms of whether they present a function as being necessary and/or sufficient for the existence of consciousness (Fig. 1c).

By locating theories in this framework, we can acknowledge which of them are offering truly competing accounts and which are compatible accounts after different goals. Researchers can then guide their investigation along this assessment of the state of the art. That is, they can proceed to determine which among the competing theories is most plausible in light of their common goal. Alternatively, they can consider the relationship between theories that address different aspects of the functions of consciousness. For example, they might seek to specify how these theories complement each other with an eye to eventually developing a unified theory of consciousness and its functions. Existing theories, however, can be ambiguous about one or more of these dimensions. In such cases, the framework helps us to clarify their character by illuminating their conceptual ambiguities and suggesting how to disambiguate.

The rest of the paper proceeds as follows. The next section provides some preliminary clarifications about function, consciousness, and existing research on the functions of consciousness. The following three sections each explain the three dimensions of the framework, (i) target, (ii) explanatory order, and (iii) necessity/sufficiency, in more detail. The final section concludes the paper by applying the three-dimensional framework to some existing theories of the functions of consciousness to demonstrate its usefulness.

Preliminary remarks: function and consciousness

Our goal is to develop a meta-theoretical framework for theories of the functions of consciousness. Before ascending to the meta-theoretical considerations, however, some preliminary remarks about the notions of consciousness and function, as well as current controversies over the functions of consciousness, are needed to specify the area of interest of this paper.

Function is an equivocal term. When we say that X bears a computational function, it roughly means that X takes certain inputs and returns certain outputs under certain background conditions (Putnam 1975). In contrast, when we say that X bears a biological function, it roughly means that X plays a certain role in sustaining an organism's life. It is standard in the philosophy of biology to distinguish between two concepts of biological functions: the "selected effect" and the "causal role" concept (Millikan 1989; Garson 2011; Laubichler *et al.* 2015; Keeling *et al.* 2019). The selected effect concept implies that a function-bearing entity X was historically selected by natural selection precisely because it bore this function. In contrast, the causal role concept only implies that X currently contributes to the operation

of the system to which it belongs. Computational and biological functions are not necessarily mutually exclusive as the latter can be re-described and explained in computational-functional terms. When we talk about the functions of consciousness in this paper, we are concerned with its biological functions yet without distinguishing the selected effect and the causal role concepts.[2]

To say that consciousness has functions does not imply that consciousness is defined by its functions. Block (1995) famously distinguishes between two notions of consciousness: access and phenomenal consciousness. Access consciousness is defined in terms of a functional role that consciousness is supposed to serve. In contrast, phenomenal consciousness is defined in terms of phenomenal properties, the property of there being something it is like to be in a mental state (Nagel 1974; Chalmers 1996). Given this distinction, one might think that debates over the functions of consciousness concern access consciousness rather than phenomenal consciousness. However, the definition of phenomenal consciousness does not imply that it plays no functional role in our cognitive life. There is no *a priori* reason to reject the possibility of there being nomological connections between phenomenal consciousness and cognitive functions; questions about the functions of consciousness are intelligible under phenomenal conceptions of consciousness. Therefore, debates about the functions of consciousness can be seen as concerning phenomenal consciousness.

Let us turn to current debates about the functions of consciousness. Various theories are proposed, drawing on different theoretical perspectives (for recent reviews, see Wiese 2020 and Black 2021). For example, Dehaene (2014) draws on experimental findings to propose that the function of consciousness is to broadcast information globally within cognitive systems. Based on the representationalist theory of consciousness, Tye (1996) argues that the function of consciousness is "to supply the conceptual centers with information from the senses, for use there in the rational control of action, and thereby to enable creatures to do a wide variety of things" (p. 301). By contrast, Kriegel (2004) appeals to phenomenological considerations and states that the functional role of consciousness is "to give the subject just enough information to know how to easily obtain fuller information about her concurrent experience" (p. 183). Birch *et al.* (2020) suggest from an evolutionary perspective that unlimited association learning is a functional marker of the transition from non-conscious to conscious beings.

Some researchers suggest that, despite the diversity of views, there are several points of agreement in this literature (Morsella 2005; Black 2021). According to Morsella (2005), for example, cognitive psychologists generally agree that consciousness functions to "integrate neural activities and information-processing structures that would otherwise be independent," which she calls the "integration consensus" (p. 1002). Black (2021) claims that there are four further points of agreement regarding the function of consciousness:

(1) Intentionality: Conscious experience has representational content that must be made accessible for further processing by the broader cognitive system.
(2) Temporality: Conscious experiences generate expectations about the future that are informed by past experience. These expectations also support the possibility of surprise, an essential aspect of learning.
(3) Adaptivity: Conscious processing facilitates learning, operant conditioning, non-automatic problem-solving, planning, and goal-oriented control over action. Unconscious processing, in contrast, facilitates automaticity, expertise, and massive parallel processing.
(4) Valence: Conscious experience assigns positive and negative valence to certain internal and external intentional contents, motivating behaviors that preserve the animal's internal homeostasis (Black 2021, 195–96).

However, some researchers refuse to accept all these points of apparent consensus (Rosenthal 2008; Lau 2009; Goff 2017). Panpsychists, such as Goff (2017), attribute a radically different role to consciousness, claiming that microphysical entities possess consciousness, which serves as a categorical basis for causal structures of the physical world, including our bodies. Epiphenomenalists claim that consciousness is nothing more than an epiphenomenon of physical processes and so plays no functional role. Some empirically minded researchers also deny that consciousness plays any significant functional role. For example, Rosenthal (2008) cites experimental research to support his argument that consciousness of non-perceptual intentional states (such as thinking and desiring) has no significant function. Likewise, Lau (2009) contends that "many sophisticated functions can be performed unconsciously or driven by unconscious information," concluding that "experiments have not yet been able to convincingly pin them down" (p. 166).

This brief overview should be enough to illustrate our previous assessment of the current theoretical landscape. There are widely varying views not only on what counts as the functions of consciousness but even on their existence. Furthermore, it is not even clear whether all these varying accounts are after the same goal when they purport to offer explanations of the functions of consciousness. Therefore, we think that a meta-theoretical approach is indispensable today to attain a clear understanding of the functions of consciousness. In the next section, we will begin our meta-theoretical consideration by introducing the first dimension, target, that constitutes our three-dimensional framework.

Target
Creature consciousness and state consciousness

Each theory of the functions of consciousness presupposes a specific notion of consciousness in determining its explanatory target—even if this is not always stated explicitly in the theory itself. Consider the distinction between "creature consciousness" and "state consciousness", which some researchers regard as crucial for studying the functions of consciousness (Dretske 1997; Rosenthal 2008) (Fig. 1a). Creature consciousness is attributed to an agent, including biological organisms and artificial cognitive systems, taken as a whole. When we say that humans and bats are conscious, but plants, embryos, and thermostats are not, we are referring to creature consciousness. It denotes the property of being conscious, which can be instantiated in various ways as agents undergo perceptual, affective, cognitive, and other forms of experience. State consciousness, in contrast, denotes the

---

[2]   The two types of biological functions can be dissociated. If X was selected for doing Y because doing Y was advantageous in the past, but X does not perform Y at the present, doing Y is a selected effect function but not a causal role function. Likewise, if X emerged as a useless byproduct of something advantageous but plays a positive role for survival and reproduction at the present, X has a causal role function but not a selected effect function. However, although some researchers aim to identify the selected effect function of consciousness (Black 2021), this is probably only possible through the identification of its causal role functions. Given that, unlike other biological traits, consciousness is not observable, it is difficult even to speculate how the possession of consciousness could have been advantageous historically without knowing how it contributes to our life at present.

property of being conscious attributed to mental states and processes, which are only components of agents taken as a whole. When we compare a conscious perceptual state, a conscious reasoning process, or conscious motor control with corresponding unconscious states and processes, we are talking about state consciousness.[3] Accordingly, theories of the functions of consciousness that explore creature consciousness and those that explore state consciousness have different explanatory targets. It follows from this that they would not necessarily be in competition even if they were making different claims about the functions of consciousness.

We propose that the difference between the functions of these two forms of consciousness corresponds to the difference between "personal-level cognitive capacity" and "sub-personal level cognitive role". How best to define the distinction is a controversial question that is still up for debate (Dennett 1969; Drayson 2014). Without getting into the details, we will assume here that personal-level cognitive capacities are properties attributed to whole agents. For instance, vision is a personal-level capacity that belongs to agents that can determine the size and shape of environmental objects based on optical information delivered through the eyes. Sub-personal-level cognitive roles are properties attributed to mental states that are components of whole agents. For instance, edge detection is a sub-personal-level role fulfilled by the visual area of the brain (or its sub-component), which is a crucial component of the whole organism that underpins its personal-level capacity for vision.

The function of creature consciousness corresponds to the personal-level capacity an organism comes to have through its possession of creature consciousness. For instance, Kanai *et al.* (2019) suggest that the function of consciousness is to endow organisms (and possibly non-organic systems) with "the ability to generate possibly counterfactual representations using internal models learned through interactions with the environment" (p. 4). As this specifies what the whole organism can do by virtue of having consciousness, we can regard it as a description of the function of creature consciousness.

In contrast, the function of state consciousness is determined by the sub-personal-level cognitive role that conscious states or processes can perform, and unconscious states or processes cannot, within a whole cognitive system. For instance, the function of a conscious visual state might be characterized in terms of its role in delivering visual information to other cognitive processes, such as motor control, planning, and imagining, within a cognitive system. For example, Dehaene (2014) argues that "the function of consciousness may be to simplify perception by drafting a summary of the current environment before voicing it out loud, in a coherent manner, to all other areas involved in memory, decision, and action" (p. 100). As this concerns the role of a conscious perceptual state within a cognitive system, we can conclude that what Dehaene is describing here is the function of perceptual state consciousness.

Our proposal is that we should clearly differentiate the functions of creature and state consciousness when exploring the functions of consciousness, because they are different in category: one at the personal level and the other at the sub-personal level. By foregrounding this conceptual distinction, we can effectively investigate the specific relations between the personal- and sub-personal-level functions of consciousness. For instance, can a personal-level function of consciousness be reduced to a set of sub-personal-level cognitive roles of corresponding conscious states? Is a sub-personal-level cognitive role of conscious states fundamentally individuated by the personal-level function of consciousness to which they contribute? Such questions can be effectively addressed only when we clearly distinguish between creature and state consciousness and thereby two distinct kinds of functions of consciousness. This is analogous to the fact that clear conceptual distinctions allow us to explore effectively the relations between, for example, phenomenal consciousness and access consciousness and the selected effect concept and causal role concept of biological function.[4]

Generic consciousness and specific consciousness

Another distinction that is important when determining the target of a theory of the functions of consciousness is that between "specific kinds" of consciousness and the "generic property" found in all kinds of consciousness (Fig. 1a). Accordingly, studies on the functions of creature consciousness can proceed in one of two possible directions: they can either study the functions of specific kinds of conscious experiences or of consciousness taken as a generic property. We will explain each of these in turn.

Creature consciousness is the property of being conscious attributed to agents taken as a whole. It is instantiated in different kinds of experiences, such as perceptual, affective, and cognitive (Kriegel 2015; Feinberg and Mallatt 2018; Shepherd 2018). The function associated with one specific kind of conscious experience would differ from that of other kinds of conscious experiences. For instance, one might think that pain has the function of motivating the subject to take certain actions in relation to body parts that are potentially in trouble (Martínez 2011; Klein 2015). Other sensory experiences, such as vision, taste, or even touch, clearly do not share this cognitive function. There may even be a kind of conscious experience so primitive that gastropods might have it (Feinberg and Mallatt 2018). Thus, the function of creature consciousness would vary depending on which kind of conscious experience is set as the explanatory target. For example, we might identify different functions for bat consciousness and human consciousness by considering the functions of their conscious experiences separately.

Despite this diversity in the kinds of conscious experiences, we can also treat creature consciousness as one generic kind. The

---

[3] Note that the distinction between creature and state consciousness does not overlap with the distinction between the level of consciousness and the content of consciousness (or the quantity of consciousness and the quality of consciousness). Creature consciousness can vary in quality. For instance, when we consider how the consciousness of human beings and the consciousness of bats differ in their content, we target creature consciousness. Likewise, state consciousness can be discussed from a quantitative perspective. For instance, when we consider whether a perceptual state is attentively conscious, peripherally conscious, or completely unconscious, in other words when we consider the conscious level of a perceptual state, we target state consciousness. The distinction between creature and state consciousness primarily concerns what consciousness is attributed to.

[4] One might refuse to utilize the distinction between creature and state consciousness for the purpose of classification, doubting that the distinction captures a real ontological difference in consciousness. The nature of the relationship between creature and state consciousness remains controversial. While McBride (1999) claims that the distinction is merely grammatical and there is no corresponding ontological difference in consciousness, Manson (2000) disputes this, stating that each is a distinct aspect of consciousness. While creature consciousness is commonly analyzed in terms of state consciousness (Bayne 2007; Brogaard *et al.* 2021), certain scholars pursue the opposite approach (Nida-Rümelin 2018). This paper, however, does not engage in these metaphysical debates. Whatever the deep ontological relationship between these two forms of consciousness is, the distinction can be justified on a practical level. It helps us to see that there are two different categories of functions of consciousness that have been explored in consciousness studies to date: personal-level capacity and sub-personal-level cognitive role. This is important in allowing further studies of the functions of consciousness to proceed without conceptual confusion.

function of this generic kind of creature consciousness can also be the target of research. Because it is shared by every kind of conscious experience, possibly including some primitive ones, it must be maximally general.

One question regarding the notion of generic creature consciousness is whether it can be identified with the "minimally conscious states" that may result from certain forms of brain damage (Giacino *et al.* 2002; Kotchoubey *et al.* 2014) or the "minimal phenomenal experience" that can emerge through deep meditation (Metzinger 2020). If either of these kinds of minimal consciousness is implicit in every form of conscious experience, even if we can hardly notice it in our usual states, we can think of it as identical to generic creature consciousness. However, if such minimal consciousness is an exceptional state that has little in common with more ordinary conscious states, we can regard it as being completely different from generic creature consciousness. At this stage, it is safer not to identify generic creature consciousness with minimal consciousness of any kind, leaving the issue open to further investigation. We can target the functions of minimal consciousness by focusing on patients with brain damage and trained meditators, without assuming that it is identical to generic creature consciousness. Similarly, we can target the function of generic creature consciousness by attempting to identify the general function common to all kinds of conscious experiences, potentially including those of primitive creatures such as insects.

State consciousness can also be instantiated variously in perceptual, affective, cognitive, and other kinds of mental states and processes. The functions of state consciousness would vary depending on what kind of conscious states and processes are being targeted. For instance, while the function of a conscious visual state may be to deliver visual information to other cognitive faculties globally, the function of a conscious reasoning process may be to disseminate amodal information across the conceptual system. As in the case of creature consciousness, we can also target a functional property that is shared by every specific conscious state and process. As both Kriegel (2004, 174) and Rosenthal (2008, 830) emphasize, we must be careful not to confuse a specific function associated only with a certain kind of conscious state or process with a common function that is shared by all kinds of conscious states and processes. A common strategy in this type of investigation is to explore the functions of specific conscious states or processes first and then abstract any generic function that is shared among them (Dehaene 2014).[5]

Note that the generic function of state consciousness might differ between different animal species (or between different types of cognitive systems). We should be careful not to equate the functions of state consciousness of human beings with those of state consciousness of other kinds of animals and of animals in general. Additional steps are needed to identify the function of the generic state consciousness of animals in general as distinct from the functions of the generic state consciousness of each kind of animal.

Explanatory order

The notion of functions of consciousness is ambiguous regarding the explanatory order between consciousness and functions. There are at least two different ways to interpret the notion (Fig. 1b). On the first interpretation, it refers to the "functional basis" of consciousness: cognitive functions that account for certain aspects of consciousness (the functional basis conception). On the second interpretation, it refers to the "functional contributions" of consciousness: cognitive functions whose presence is explained in terms of certain aspects of consciousness (the functional contribution conception).

Let us formulate each conception more formally. Below, we stipulate that *F* means function, *C* means consciousness, and "in virtue of" expresses the order of explanation. The functional basis conception then holds that *C* is in place in virtue of *F*. When this explanatory order holds between *C* and *F*, *F* is the functional basis of *C*. The functional contribution conception holds that *F* is in place in virtue of *C*. When this explanatory order holds between *F* and *C*, *F* is the functional contribution of *C*.

These two conceptions, the functional basis and the functional contribution, are incompatible if *C* and *F* are fully fixed. In that case, if we explain *C* in virtue of *F*, then we cannot explain *F* in virtue of *C* on pain of vicious circularity, and vice versa. However, we can properly apply these two conceptions to different kinds of consciousness, such as creature consciousness and state consciousness, simultaneously. For example, we can apply the functional basis conception to state consciousness by stating that a mental state becomes conscious in virtue of performing a certain cognitive role; at the same time, we can apply the functional contribution conception to creature consciousness by stating that a creature gains a certain capacity in virtue of having consciousness.

This distinction between functional basis and functional contribution is theoretically important as some questions are only answerable based on either conception of the functions of consciousness. As an example, we can consider the question of the evolutionary origin of consciousness. In exploring this question, one may start by arguing or assuming that consciousness survived natural selection because it played some adaptive function in evolutionary history (Nichols and Grantham 2000; Feinberg and Mallatt 2016). In this case, the adaptive function is considered to be in place in virtue of the presence of consciousness, which means that this approach implicitly presupposes the functional contribution conception.

By contrast, we cannot sensibly approach the question of the evolutionary origin of consciousness if we base our exploration on the functional basis conception. According to this conception, *F* is the function of mechanisms that account for *C* rather than the function that *C* performs. However, it is the latter that may directly explain why consciousness survived natural selection. The functional basis conception is even compatible with epiphenomenalism, according to which consciousness is just an epiphenomenal by-product of its functional basis and lacks any adaptive function of its own. If epiphenomenalism is true, consciousness as such cannot be selected through evolutionary history in virtue of its adaptive function because it does not have one in the first place. Certainly, even if epiphenomenalism is true, we can indirectly explain why consciousness was selected if we can reveal how the functional basis itself contributed to the survival of its bearer. However, this does not constitute an explanation of why consciousness was selected in terms of its own functional properties. Accordingly, if we want to understand

---

5   Once we know the generic function of state consciousness, it might be possible to deduce the functions of specific types of conscious states. Suppose that the generic function of state consciousness is to globally distribute information. Then we might infer that the function of conscious perceptual state is the global distribution of perceptual information, the function of conscious memory is the global distribution of remembered information, and so on. Whether we can derive all specific functions of state consciousness based on this deductive method is an open question requiring further empirical research.

why consciousness survived natural selection in terms of its own properties, we should base our investigation on the functional contribution conception of the function of consciousness.

If our aim is to answer the question of how consciousness occurs in a physical system by exploring the functions of consciousness, we should base our approach on the functional basis conception: *C* is in place in virtue of *F*. This conception allows us to explain how *C* occurs in a physical system by specifying *F* and uncovering how *F* is realized by the physical architecture. However, this strategy does not work if we adopt the functional contribution conception, according to which *F* is in place in virtue of *C*. Even if we were to successfully specify the functional contribution of consciousness, we could not explain how consciousness occurs in a physical system by uncovering how the functional contribution is realized by the physical architecture. This is because, given that functional contribution is grounded in the presence of consciousness, we cannot specify how a functional contribution of consciousness is realized by a physical system without knowing how consciousness is realized by the physical system in the first place, which just brings us back to the original question. Thus, the identification of the functional contribution of consciousness does not help us to address the question of how consciousness occurs in a physical system.

One might suspect that the distinction between the functional basis conception and the functional contribution conception would collapse if we adopt the functional identity view of consciousness, according to which *C* is identical to *F*. Because the identity relation is symmetrical and the in-virtue-of relation is asymmetrical, it appears that if the identity relation holds between *C* and *F*, the former cannot explain the latter and vice versa. Additionally, the functional identity view of consciousness allows us to systematically address the questions of the evolutionary origin of consciousness and the physical basis of consciousness. That is to say, we can systematically address why consciousness was selected by natural selection through exploration of how the function contributed to its bearers, and how consciousness is realized by a physical system through exploration of how the function is implemented by the physical system. This is a simple and parsimonious research paradigm. Thus, it may be suggested, if the functional identity view is true, the distinction between functional basis and functional contribution has no theoretical importance.

To address this suspicion, we first emphasize that we can apply different conceptions of the relation between consciousness and function to different kinds of consciousness. As shown above, for instance, we can apply the functional basis conception and the functional contribution conception separately to creature consciousness and state consciousness. Given this, it is possible to apply the functional identity view to state consciousness, asserting that conscious mental states are identical to their functional roles, while applying the functional contribution conception to creature consciousness. Likewise, it is possible to adopt the functional identity view for creature consciousness, asserting that every conscious experience is identical to its corresponding cognitive capacity, while applying the functional basis conception to state consciousness. Thus, even if we apply a functional identity view to some kinds of consciousness, there will be room to apply another conception of function to other kinds of consciousness.

One might wonder what to make of the total functional identity view, according to which every kind of consciousness is identical to its function. If this view is true, the distinction between the functional basis conception and the functional contribution conception plays no theoretical role. However, philosophers have raised many objections to the total functional identity view. For instance, Kriegel (2004) argues against the total functional identity view by focusing on the distinction between a functional property and its categorical basis, claiming that phenomenal consciousness should be counted as not a function but its categorical basis. Chalmers (2010) argued against the total functional identity view by stating that the reductive explanation of phenomenal consciousness in functional terms does not work. It is at least fair to say that we should not naively adopt the total functional identity view without being able to defend it from such objections. We may be able to plausibly defend a partial functional identity view stating that a specific kind of consciousness is identical to its function (Block 1995), but it is much more difficult to defend the total version. Given this, we should consider the functional basis conception and functional contribution conception as possible distinct options for defining the relation between consciousness and function, without assuming the total functional identity view.[6]

Despite the theoretical importance of the dimension of explanatory order, many theorists of the functions of consciousness do not explicitly take a clear stance on it; indeed, they sometimes seem to move back and forth between different conceptions. For instance, Kanai *et al.* (2019) title their paper "Information generation as a functional basis of consciousness," which suggests that they are concerned with the functional basis of consciousness. However, they open their abstract by asking "what are the biological advantages of having consciousness?," which suggests that they are concerned with the functional contribution of consciousness. Likewise, Chapter 3 of Dehaene (2014) discusses what consciousness is good for, which clearly concerns the functional contribution of consciousness. However, he also writes "[t]o be consciously accessible, information must be encoded as an organized pattern of neuronal activity in higher cortical regions, and this pattern must, in turn, ignite an inner circle of tightly interconnected areas forming a global workspace" (p. 198). This quotation suggests that he is addressing the functional basis of consciousness because it seems to describe the functional profile that makes informational states conscious. Furthermore, he states that "I believe that consciousness reduces to what the workspace does" (p. 168) and that "consciousness is global information sharing," (p. 163), which suggests that he is adopting a functional identity view.

This shows that such theories can benefit from further clarification. If different explanatory relations seem to be posited within a single theory of the functions of consciousness, we should demand that the theory clarifies what *C* refers to for each explanatory relation. This clarification will help to conceptually systemize each theory of the functions of consciousness, including those of Kanai *et al.* (2019) and Dehaene (2014), illuminating the tasks for each theory to tackle. For instance, we might be able to interpret Dehaene (2014) as taking the functional basis conception for state consciousness and the functional contribution conception for creature consciousness. That is to say, it might be possible to interpret the theory as claiming that a mental state or process

---

6   Furthermore, it is not entirely clear whether the identity relation is incompatible with the in-virtue-of relation (Azzano and Carrara 2021). For instance, many would accept both that H_2O is ontologically identical to water and that various properties of water are explained by the fact that it is H_2O. If this holds for the relation between consciousness and its function, even taking the total functional identity view may not conflict with applying either the functional basis conception or the functional contribution conception at the same time. This is why we did not list the functional identity view along with the functional basis conception and the functional contribution conception on the dimension of explanatory order.

becomes conscious in virtue of performing a certain role within a whole cognitive system and that the whole cognitive system gains a certain capacity in virtue of having consciousness. If this is correct, the next task for this theory will be to clarify (i) the relationship between state consciousness and creature consciousness and (ii) its conceptual relationship with the functional identity view: for example, whether it entails this view, is compatible with but not strictly wedded to it, or is incompatible with it.

Necessity/sufficiency

The third dimension concerns the necessity/sufficiency relation between consciousness (C) and function (F). We can distinguish between necessary and sufficient conditions holding between C and F (where necessity and sufficiency are understood as nomological rather than metaphysical). By combining this distinction with the one about explanatory order, we arrive at the following four positions (where C denotes either creature consciousness or state consciousness):

(i) The necessary functional basis conception (F_{bn}): (1) C is in place in virtue of F (and perhaps some other conditions) and (2) C cannot be in place without F (F is necessary for C).

(ii) The sufficient functional basis conception (F_{bs}): (1) C is in place in virtue of F and (2) C is always in place if F occurs (F is sufficient for C).

(iii) The sufficient functional contribution conception (F_{cs}): (1) F is in place in virtue of C (and perhaps some other conditions) and (2) F cannot be in place without C (F is sufficient for C).

(iv) The necessary functional contribution conception (F_{cn}): (1) F is in place in virtue of C and (2) F is always in place if C occurs (F is necessary for C).

We notate F in (i) to (iv) as, respectively, F_{bn}, F_{bs}, F_{cs}, and F_{cn}. F_{bn} and F_{cn} are functions necessary for consciousness, and F_{bs} and F_{cs} are functions sufficient for consciousness. We can then clarify theories of the functions of consciousness in terms of whether they aim to deal with the functions necessary for consciousness (F_{bn} or F_{cn}) or the functions sufficient for consciousness (F_{bs} or F_{cs}) (Fig. 1c).

It is important to note that many functions can be necessary for consciousness but *uninformative*. Consider, for example, the property of being capable of processing information. It is almost a truism that the general capacity to process information is necessary for the existence of creature consciousness. Thus, the function of processing information can be regarded as either an F_{bn} or an F_{cn}. However, the identification of the function of processing information as F_{bn} or F_{cn} does not reveal anything novel; it does not advance our understanding of the functions of consciousness. In this sense, it is uninformative to specify the function of processing information as F_{bn} or F_{cn}.

To make the identification of the Fs necessary for C more informative, the descriptive content should be as extensive as possible and ideally maximal. The descriptive content of F necessary for C is maximal if F fails to be necessary for consciousness when any other content is added to it. The maximal description is informatively ideal for the identification of F necessary for C. For instance, we can add more specific conditions for information processing that are necessary for C and can thereby obtain a more informative account of the F necessary for C. This maximization process is required to achieve better identification of any F necessary for C.

Similarly, there can be many functions that are sufficient for consciousness, the specification of which is uninformative, such as the property of being capable of doing everything that David Chalmers can do (call it the property of 'chalmerizing'). It is almost a truism that having the property of chalmerizing is (nomologically) sufficient for the existence of creature consciousness.[7] It is uninformative, however, to identify the property of chalmerizing as either F_{bs} or F_{cs} because that identification does not advance our knowledge of the functions of consciousness. In order to make the identification of a function sufficient for consciousness more informative, its descriptive content should be as less extensive as possible and ideally minimal. The descriptive content of F sufficient for C is minimal if F fails to be sufficient if any part of the descriptive content is removed. The minimal description is informatively ideal for the identification of F sufficient for C. For instance, one may attempt to reduce the functional components of chalmerizing to make its descriptive content less extensive and ideally minimal. This minimization process is required to achieve better identification of F sufficient for C.

In general, we seek informative identifications of the functions of consciousness. When exploring the functions necessary for consciousness, we should seek maximal descriptions. In contrast, when exploring the functions sufficient for consciousness, we should seek minimal descriptions. Indeed, many theorists of the functions of consciousness abide by this principle in their investigations. For instance, Wiese (2020) argues that many existing theories of the functions of consciousness share the assumption that "information generation" is the functional property necessary for consciousness. As part of this argument, Wiese attempts to characterize information generation in various terms, such as "an active process of producing representations that are based on stored information and have been adapted to current purposes" (p. 3) and "[information generation] seems to require that the information be stored in a way that enables using it for different purposes (e.g. by compressing it), which means that further processing is necessary before the stored information can be used" (p. 3). This can be interpreted as an attempt to specify a function necessary for consciousness as maximally as possible. Furthermore, Koch *et al.* (2016) survey many attempts to identify the neural correlates of consciousness (NCC), which they characterize as "the minimum neuronal mechanisms jointly sufficient for" consciousness (p. 308). If we try to infer cognitive functions from NCC (Seth 2009), such cognitive functions are supposed to be the minimal functions sufficient for consciousness.

Note that the maximal function necessary for consciousness does not necessarily overlap with the minimal function sufficient for consciousness. Firstly, the maximal function necessary for consciousness may not be sufficient for consciousness. There may be a further disjunctive set of enabling conditions, one of which needs to be satisfied for consciousness to occur but none of which is necessary for consciousness (what we call the possibility of disjunctive enabling conditions). In this case, if any one of the disjunctive enabling conditions is not satisfied, consciousness does not occur even though the maximal function necessary for consciousness is realized. Secondly, the minimal function sufficient for consciousness may not be necessary for consciousness. Perhaps there are some ways to functionally realize consciousness

---

7 On the functional basis conception (F_{bs}), the truism is analyzed as follows: the creature has consciousness necessarily in virtue of having the property of chalmerizing. In contrast, on functional contribution conception (F_{cs}), the truism is analyzed as follows: having consciousness contributes to instantiating the property of chalmerizing, and a creature cannot have the property of chalmerizing without having consciousness.

(namely there are different sets of minimal functions sufficient for consciousness) such that no functional element is shared among all instances of consciousness (what we call the possibility of multiple realizability). Whether there is a specific function necessary and sufficient for consciousness remains an open question.

This question needs to be addressed to argue for the functional identity view, according to which C is identical to F. This view presupposes that there is a function that is necessary and sufficient for consciousness because the identity relation between X and Y implies that X is necessary and sufficient for Y. Proponents of the functional identity view thus need to argue against both possibilities mentioned above, namely disjunctive enabling conditions and radical multiple realizations.

Conclusion:the three-dimensional framework and its usefulness

In this paper, we have proposed a three-dimensional framework to study the functions of consciousness (Fig. 1). On our account, theories in this area can be classified in terms of:

(i) Target: Which notion of consciousness they presuppose in determining their explanatory target: creature consciousness or state consciousness, and specific or generic kinds of consciousness.

(ii) Explanatory order: Which explanatory relation between consciousness and function they posit, namely, the functional basis conception or the functional contribution conception.

(iii) Necessity/sufficiency: Which relation between consciousness and function they identify, namely, a (maximal) function necessary for consciousness or a (minimal) set of functions sufficient for consciousness.

This three-dimensional framework can also be used to clarify existing theories of the functions of consciousness. We can identify the theoretical ambiguities of such theories within our framework. For instance, we saw that Dehaene (2014) and Kanai *et al.* (2019) are ambiguous on the dimension of ("Explanatory order"). This finding puts pressure on the theorists to disambiguate what kind of consciousness their theory targets and which explanatory relation between consciousness and function it presupposes. This process helps us to see the theoretical challenges their theories need to tackle. This function of clarification is an intended role of our framework.

While some theories of the functions of consciousness are ambiguous regarding one or more dimensions, others fall squarely into our framework. For example, Morsella's (2005) supramodular interaction theory holds that conscious states "allow information from diverse sources to interact in order to produce adaptive action" and "are required for the outputs of different supramodular response systems to interact" (p. 1012). This suggests that the supramodular interaction theory (i) targets state consciousness, (ii) explores the functional contribution of state consciousness, and (iii) seeks to identify functions that are sufficient for the existence of conscious states.

By locating such existing theories in our framework, we can also clarify the relationship between them. For instance, Rosenthal (2008) states that there is no significant function of consciousness. In contrast, Kriegel (2004) claims that the function of consciousness is to give the subject just enough information to know how to easily obtain fuller information about her concurrent experience. On the face of it, these two claims appear to be in direct conflict. However, Rosenthal's (2008) theory

(i) targets the function of state consciousness specifically of non-sensory kinds, (ii) explores the functional contribution of state consciousness, and (iii) seeks to identify functions that are sufficient for the existence of conscious states. By contrast, Kriegel's (2004) theory (i) targets the function of creature consciousness of a generic kind, (ii) explores the functional contribution of creature consciousness, and (iii) seeks to identify functions that are sufficient for the existence of creature consciousness.[8] In light of this, we can see that Rosenthal (2008) and Kriegel (2004) are not in direct conflict, because the former addresses non-sensory kinds of state consciousness and the corresponding sub-personal-level cognitive role, and the latter addresses creature consciousness and the corresponding personal-level capacity.

Let us briefly discuss how the integrated information theory (IIT) of consciousness can be located within our framework. We interpret IIT as (i) targeting state consciousness, (ii) exploring its functional basis, and (iii) attempting to identify its functional role that is necessary and sufficient for it. According to IIT, a physical complex bears consciousness when it has the locally highest value of integrated information, which represents the degree of differentiation and integration among its parts (Oizumi *et al.* 2014; Tononi *et al.* 2016; Miyahara and Witkowski 2019; Cea 2021). Consciousness is supposed to be explained by the functional role of a system to integrate information that operates at the sub-personal level (Tononi 2004). In other words, IIT targets the function of state consciousness, presupposing the functional basis conception.[9] In addition, the "central identity" of IIT implies that integrated information is necessary and sufficient for consciousness (Tononi and Koch 2015).

In conclusion, our three-dimensional framework can pave the path toward a unified theory of the functions of consciousness by making it possible to clarify, classify, and compare existing theories in a systematic way. Although it is beyond the scope of our paper to achieve this for every theory of the functions of consciousness, we have illustrated the idea by applying the framework to some theories, such as the phenomenological approach (Kriegel 2004), global workspace theory (Dehaene 2014), information generation theory (Kanai *et al.* 2019) and IIT (Oizumi *et al.* 2014; Tononi *et al.* 2016). Further research is needed to clarify and classify other theories of consciousness, for example, the attentional schema theory (Graziano and Webb 2015; Graziano *et al.* 2020) and predictive error minimization theory (Hohwy 2015).

This framework is open to further development. For instance, we might be able to add beliefs about consciousness in the target dimension. In fact, some explore the function of cognitive systems that enable us to have beliefs about consciousness (Clark 2019). In contrast, some attempt to identify the functions of having the belief that there exists phenomenal consciousness (Frankish 2016; Kammerer 2019). When properly updated, our framework could also situate these views in the theoretical landscape concerning the function of consciousness.

---

8   Although Kriegel's (2004) theory may appear to address state consciousness (he often uses the term "conscious state"), his argument is based on the phenomenological approach, which explores the ways in which an agent's (or a creature's) conscious experiences contribute to its cognitive life. Thus, Kriegel's (2004) theory should be interpreted as concerning the personal-level function of creature consciousness rather than the sub-personal-level function of state consciousness.

9   Cea (2021) advances an alternative interpretation of IIT. She argues that IIT implies that consciousness is the "functional-realizer" of information integration. In this case, IIT presupposes the functional contribution conception.

Finally, our meta-theoretical approach to the function of consciousness can be integrated with other meta-theoretical studies that aim to compare and classify general theories of consciousness (Niikawa 2020; Doerig *et al.* 2021; Pin *et al.* 2021; Sattin *et al.* 2021; Signorelli *et al.* 2021). This integration could provide a more comprehensive and detailed framework in which to locate theories of consciousness.

Data availability

No data are associated with this article.

Funding

Conflict of interest statement

None declared.

References

Azzano L, Carrara M. The grounding of identities. *Philosophia* 2021;**49**:1943–52.

Bayne T. Conscious states and conscious creatures: explanation in the scientific study of consciousness. *Philos Perspect Suppl Nous* 2007;**21**:1–22.

Birch J, Ginsburg S, Jablonka E. Unlimited associative learning and the origins of consciousness: a primer and some predictions. *Biol Philos* 2020;**35**:56.

Black D. Analyzing the etiological functions of consciousness. *Phenomenol Cogn Sci* 2021;**20**:191–216.

Block N. On a confusion about a function of consciousness. *Behav Brain Sci* 1995;**18**:227–47.

Brogaard B, Chomanski B, Gatzia DE. Consciousness and information integration. *Synthese* 2021;**198**:763–92.

Cea I. Integrated information theory of consciousness is a functionalist emergentism. *Synthese* 2021;**199**:2199–224.

Chalmers DJ. *The Conscious Mind: In Search of a Fundamental Theory*. New York: Oxford University Press, 1996.

Chalmers DJ. *The Character of Consciousness*. New York: Oxford University Press, 2010.

Clark A. Consciousness as generative entanglement. *J Philos* 2019;**116**:645–62.

Cohen MA, Dennett DC. Consciousness cannot be separated from function. *Trends Cogn Sci* 2011;**15**:358–64.

Dehaene S. *Consciousness and the Brain: Deciphering How the Brain Codes Our Thoughts*. New York: Penguin Books, 2014.

Dennett DC. *Content and Consciousness*. Oxford: Humanities Press, 1969.

Dennett DC. Facing up to the hard question of consciousness. *Philos Trans R Soc Lond B Biol Sci* 2018;**373**: 20170342.

Doerig A, Schurger A, Herzog MH. Hard criteria for empirical theories of consciousness. *Cogn Neurosci* 2021;**12**:41–62.

Drayson Z. The personal/subpersonal distinction. *Philos Compass* 2014;**9**:338–46.

Dretske F. What good is consciousness? *Can J Philos* 1997;**27**:1–15.

Feinberg TE, Mallatt JM. *The Ancient Origins of Consciousness: How the Brain Created Experience*. Cambridge: The MIT Press, 2016.

Feinberg TE, Mallatt JM. *Consciousness Demystified*. Cambridge: The MIT Press, 2018.

Frankish K. Illusionism as a theory of consciousness. *J Conscious Stud* 2016;**23**:11–39.

Garson J. Selected effects and causal role functions in the brain: the case for an etiological approach to neuroscience. *Biol Philos* 2011;**26**:547–65.

Giacino JT, Ashwal S, Childs N *et al.* The minimally conscious state: definition and diagnostic criteria. *Neurology* 2002;**58**:349–53.

Ginsburg S, Jablonka E. *The Evolution of the Sensitive Soul: Learning and the Origins of Consciousness*. Cambridge: The MIT Press, 2019.

Goff P. *Consciousness and Fundamental Reality*. New York: Oxford University Press, 2017.

Graziano MSA, Guterstam A, Bio BJ *et al.* Toward a standard model of consciousness: reconciling the attention schema, global workspace, higher-order thought, and illusionist theories. *Cogn Neuropsychol* 2020;**37**:155–72.

Graziano MSA, Webb TW. The attention schema theory: a mechanistic account of subjective awareness. *Front Psychol* 2015; **6**:500.

Hohwy J. Prediction error minimization, mental and developmental disorder, and statistical theories of consciousness. In: Gennaro RJ (ed.), *Disturbed Consciousness: New Essays on Psychopathology and Theories of Consciousness*. Cambridge: The MIT Press, 2015, 293–324.

Kammerer F. The meta-problem of consciousness and the evidential approach. *J Conscious Stud* 2019;**26**:124–35.

Kanai R, Chang A, Yen Y *et al.* Information generation as a functional basis of consciousness. *Neurosci Conscious* 2019;**2019**:niz016.

Keeling DM, Garza P, Nartey CM *et al.* Philosophy of biology: The meanings of 'function' in biology and the problematic case of de novo gene emergence. *eLife* 2019;**8**: e47014.

Klein C. What pain asymbolia really shows. *Mind* 2015;**124**: 493–516.

Koch C, Massimini M, Boly M *et al.* Neural correlates of consciousness: progress and problems. *Nat Rev Neurosci* 2016;**17**:307–21.

Kotchoubey B, Vogel D, Lang S *et al.* What kind of consciousness is minimal? *Brain Inj* 2014;**28**:1156–63.

Kriegel U. The functional role of consciousness: a phenomenological approach. *Phenomenol Cogn Sci* 2004;**3**:171–93.

Kriegel U. *The Varieties of Consciousness*. New York: Oxford University Press, 2015.

Lau H. Volition and the function of consciousness. *Faith Philos* 2009;**26**:537–52.

Laubichler MD, Stadler PF, Prohaska SJ *et al.* The relativity of biological function. *Theory Biosci* 2015;**134**:143–7.

Manson NC. State consciousness and creature consciousness: a real distinction. *Philos Psychol* 2000;**13**:405–10.

Martínez M. Imperative content and the painfulness of pain. *Phenomenol Cogn Sci* 2011;**10**:67–90.

McBride R. Consciousness and the state/transitive/creature distinction. *Philos Psychol* 1999;**12**:181–96.

Metzinger T. Minimal phenomenal experience. *Philos Mind Sci* 2020;**1**:1–44.

Millikan RG. An ambiguity in the notion 'function'. *Biol Philos* 1989;**4**:172–6.

Miyahara K, Witkowski O. The integrated structure of consciousness: phenomenal content, subjective attitude, and noetic complex. *Phenomenol Cogn Sci* 2019;**18**:731–58.

Morsella E. The function of phenomenal states: supramodular interaction theory. *Psychol Rev* 2005;**112**:1000–21.

Morsella E, Andrew Poehlman T. The inevitable contrast: conscious vs. unconscious processes in action control. *Front Psychol* 2013;**4**:590.

Nagel T. What is it like to be a bat? *Philos Rev* 1974;**83**:435–50.

Nichols S, Grantham T. Adaptive complexity and phenomenal consciousness. *Philos Sci* 2000;**67**:648–70.

Nida-Rümelin M. The experience property frame work: a misleading paradigm. *Synthese* 2018;**195**:3361–87.

Niikawa T. A map of consciousness studies: questions and approaches. *Front Psychol* 2020;**11**:530152.

Oizumi M, Albantakis L, Tononi G. From the phenomenology to the mechanisms of consciousness: integrated information theory 3.0. *PLoS Comput Biol* 2014;**10**:e1003588.

Pierson LM, Trout M. What is consciousness for? *New Ideas Psychol* 2017;**47**:62–71.

Pin D, Hviid S, Skóra Z *et al.* Comparing theories of consciousness: why it matters and how to do it. *Neurosci Conscious* 2021;**2021**:niab019.

Putnam H. The nature of mental states. In: Putnam H (ed.), *Philosophical Papers: Volume 2: Mind, Language and Reality*. Cambridge: Cambridge University Press, 1975, 429–40.

Rosenthal DM. Consciousness and its function. *Neuropsychologia* 2008;**46**:829–40.

Sattin D, Magnani FG, Bartesaghi L *et al.* Theoretical models of consciousness: a scoping review. *Brain Sci* 2021;**11**:535.

Seth A. Explanatory correlates of consciousness: theoretical and computational challenges. *Cognit Comput* 2009;**1**: 50–63.

Shepherd J. *Consciousness and Moral Status*. New York: Routledge, 2018.

Signorelli CM, Szczotka J, Prentner R. Explanatory profiles of models of consciousness - towards a systematic classification. *Neurosci Conscious* 2021;**2021**:niab021.

Tononi G. An information integration theory of consciousness. *BMC Neurosci* 2004;**5**:42.

Tononi G, Boly M, Massimini M *et al.* Integrated information theory: from consciousness to its physical substrate. *Nat Rev Neurosci* 2016;**17**:450–61.

Tononi G, Koch C. Consciousness: here, there and everywhere? *Philos Trans R Soc Lond B Biol Sci* 2015;**370**: 20140167.

Tye M. The function of consciousness. *Noûs* 1996;**30**:287–305.

Wiese W. The science of consciousness does not need another theory, it needs a minimal unifying model. *Neurosci Conscious* 2020;**2020**:niaa013.