


ARTICLE

The Problem of Hard and Easy Problems

Tudor M. Baetu 

Université du Québec à Trois-Rivières, Département de philosophie et des arts, 3351, boul. des Forges, Trois-Rivières (Québec)
G8Z 4M3, Canada
Email: tudor-mihai.baetu@uqtr.ca

Abstract

David Chalmers advocates the view that the phenomenon of consciousness is fundamentally different from all other phenomena studied in the life sciences, positing a uniquely hard problem that precludes the possibility of a mechanistic explanation. In this paper, I evaluate three demarcation criteria for dividing phenomena into hard and easy problems: functional definability, the puzzle of the accompanying phenomenon, and the first-person data of subjective experience. I argue that none of the proposed criteria can accurately discriminate between the phenomenon of consciousness and mechanistically explainable phenomena.

Keywords: Consciousness; the hard problem of consciousness; philosophy of biology; mechanism; causal inference; criterion accuracy

1. Introduction

“Consciousness is not just business as usual,” David Chalmers (1996, x) assures us. There is something intrinsically and fundamentally special about the phenomenon of consciousness which posits a hard problem unlike any other in science. Judging by the numerous references to Chalmers’s claim, it would seem that most scientists in the field are comfortable with this verdict. Be that as it may, fleshing out a clear and coherent account of how the phenomenon of consciousness is intrinsically and fundamentally different from other phenomena is not as trivial as it may seem. The claim is not that consciousness is of special interest to us, that its study was shunned for historical and ideological reasons, or that some express the conviction that consciousness is immaterial and unexplainable. Such differences are extrinsic since they concern attitudes and beliefs about the phenomenon, not the phenomenon itself. Nor is it a question here of the trivially true fact that every phenomenon is special in the sense that it is different from all other phenomena (it is measured by different techniques, replicated in different experimental models, etc.). What Chalmers claims is that there is something about the phenomenon of consciousness which renders it a priori incompatible with currently accepted canons of scientific explanation.

What, then, makes the phenomenon of consciousness so special? For the past three decades, Chalmers championed the notion that conceptual analysis suffices to identify criteria for dividing phenomena into ‘easy’ and ‘hard problems.’ The details changed over the years, as Chalmers updated his own views about what constitutes scientific “business as usual,” but the demarcation line between easy and hard problems remained unchanged, invariably yielding the same verdict that consciousness constitutes a uniquely intractable hard problem, while most, if not all other biological and psychological phenomena can be construed as easy problems science can explain.

In this paper, I focus on three demarcation criteria presented in his 2010 book, *The Character of Consciousness*. According to a first criterion, the distinction between hard and easy problems hinges on the notion of functional definability, which cuts across the empirical reality of the life sciences,

dividing it into a mechanistically unexplainable phenomenon of consciousness and the rest of biological and psychological phenomena, all or most of which can be explained mechanistically. The argument here is that if a phenomenon is functionally definable, then all it could possibly take to explain it is the specification of a mechanism. However, since consciousness is not about functions, it is not amenable to a mechanistic explanation. A second criterion stipulates that we can legitimately ask why the performance of certain cognitive and behavioral functions is accompanied by subjective experience. In contrast, asking a similar question in respect to a biological concept makes no sense. Finally, according to a third criterion, mechanistically explainable objective functioning can only explain objective third-person data. But consciousness is characterized by subjective first-person data, which are not about objective functioning. Hence, mechanistic explanations of objective functions leave subjective experience unexplained.

The goal of this paper is not to defend intuitions about the mechanistic explainability or unexplainability of consciousness. Rather, the goal is methodological: to evaluate the extent to which the above classification criteria succeed in discriminating consciousness from mechanistically explained phenomena. If, by applying these criteria, one ends up placing consciousness in one category and mechanistically explained phenomena in a different category, then this result can justify Chalmers's contention that the phenomenon of consciousness posits a uniquely hard problem. If, on the other hand, the criteria fail to generate the desired classification, then Chalmers's intuitions about how consciousness is radically different from other phenomena are mistaken.

Thus, the question addressed in this paper is, 'Do the above-listed criteria work as intended by Chalmers?' Addressing this question requires an evaluation of the actual-world sensitivity and specificity of the criteria—that is, of the extent to which the application of the criteria generate false negatives, placing the hard problem of consciousness on the side of mechanistically explainable phenomena, and false positives, classifying easy problems as mechanistically unexplainable.¹ The evaluation conducted in the paper supports the conclusion that none of the three proposed criteria can accurately discriminate between the hard problem of consciousness and the easy problems of mechanistically explainable phenomena. In other words, a user will not succeed in classifying consciousness as a unique/almost unique mechanistically unexplainable phenomenon based on the presence or absence of the markers probed by Chalmers's three criteria. Of course, this doesn't prove that consciousness is not a mechanistically intractable hard problem. However, the fact that the criteria fail to work as advertised indicates that a priori intuitions about what is and isn't mechanistically explainable are unreliable.

2. The 'functional indefinability' criterion

2.a *Functional definability and mechanistic explainability*

Chalmers famously distinguishes between two explanatory projects within a science of consciousness, the hard and the easy problems:

The easy problems are easy precisely because they concern the explanation of cognitive abilities and functions. To explain a cognitive function, we need only specify a mechanism that can perform the function. [...] By contrast, the hard problem is hard precisely because it is not a problem about the performance of functions. The problem persists even when the performance of all of the relevant functions is explained. (2010, 6)

The relationship between functions and mechanisms is taken to be a conceptual truism: "How do we explain the performance of a function? By specifying a mechanism that performs the

¹Ideally, one would conduct a quantitative statistical analysis on a random sample of phenomena. However, since this is a philosophy paper, I will rely on a qualitative analysis which can only determine whether the proposed criteria have the perfect or near-perfect accuracy Chalmers attributes them.

function” (2010, 7). For instance, “[a]ll it could *possibly* take to explain reportability is an explanation of how the relevant function is performed” (7), where the ‘how’ in question is “a story about the organization of the physical system that allows it to react to environmental stimulation and produce behavior in the appropriate sorts of ways” (1996, 22). Thus, if a phenomenon is functionally definable, it follows that a mechanism can explain that phenomenon. In contrast, consciousness is not about the performance of functions, which entails that there is no ‘functional how’ to be explained mechanistically in the first place. Thus, the criterion of functional definability discriminates between consciousness and mechanistic unexplainability on one side, and all/most other phenomena and mechanistic explainability on the other.

2.b Functions and mechanisms

Before we can test the accuracy of the criterion, we need a more precise characterization of the terms ‘mechanism’ and ‘function.’ The new mechanistic philosophy offers close to a dozen characterizations of mechanisms (Glennan 2017). Among these, the one that best matches Chalmers’s terminology is the proposal that a “mechanism is a structure performing a function in virtue of its component parts, component operations, and their organization. The orchestrated functioning of the mechanism is responsible for one or more phenomena” (Bechtel and Abrahamsen 2005, 423). The mechanism is a physical entity consisting of an organized system of interacting parts, while the phenomenon for which it is responsible is the behaviour of the system, typically described as an input-output or stimulus-response sequence (Machamer et al. 2000).

Chalmers defines the term ‘function’ as “any causal role in the production of behaviour that a system might perform” (2010, 6). This understanding of the term is distinct from the systemic and evolutionary concepts of function developed in the philosophy of biology. The former requires that functional ascriptions are made in light of an explanatory account—usually, but not necessarily, of a mechanistic variety—specifying how a part of a system causally contributes to the ability of the system to behave in a certain way (Craver 2001; Cummins 1975). The latter defines functions relative to a natural selection mechanism (Millikan 1989; Neander 1991). Chalmers’s characterization best matches a broader notion of function commonly found in experimental biology, where the term is used to describe the results of controlled experiments designed to demonstrate causal relevance. Something is said to have a function if that something (mechanism, mechanistic component, factor, independent variable) makes a difference vis-à-vis something else (phenomenon, outcome, dependent variable) in the context of an experimental setup (animal or cell model, in vitro reconstitution system, etc.). Functional ascription is made here solely in virtue of evidence for causal relevance, without the explicit involvement of an explanatory account. For instance, when early geneticists and developmental biologists concluded that the function of the *wingless* gene is to determine wing development and body axis formation, they were merely describing differences between phenotypes of *Drosophila* with (test/mutated) and without mutations (wild-type/control) at the *wingless* locus. Such differences reveal that *wingless* is causally relevant to, and in this sense can be attributed the function of contributing to the outcomes of wing development and body axis formation (Sharma 1973).

It is possible, however, that Chalmers contemplates a somewhat more restrictive notion of function, of the sort assumed in functionalist theories of the mind (McLaughlin 2006). The latter define mental states in terms of causal relations to stimuli, other mental states, and behavioral responses. Here, a function is not just any cause contributing to an outcome, but more specifically a causal mediator between stimuli and responses. Many functional ascriptions in experimental science aggregate these two slightly different notions of function. For instance, since lesions to the visual cortex area V5 lead to an inability to perceive motion when presented with visual motion stimuli, while the artificial stimulation of V5 neurons affects the perception of motion (as contrasted to normal/control subjects), it is common practice to summarize causal inferences in statements such as “the principal function of V5 is to detect and signal the presence and direction of visual motion” (Zeki 2015).

2.c The sensitivity of the functional undefinability criterion remains unproven

Equipped with these precisions, we can proceed to the evaluation of the first axis of Chalmers's discrimination criterion, namely the claim that unlike other biological and psychological phenomena, consciousness is not functionally definable. The first question that comes to mind concerns the sensitivity of the criterion: Is it indeed the case that consciousness is functionally undefinable?

The empirical justification of the criterion rests on the fact that in some brain lesion patients or under experimental conditions, certain perceptual and cognitive tasks can be performed nonconsciously, thus demonstrating that task performance and behavioral responses to stimuli can be dissociated from conscious experience. Known examples include unconscious processing of visual stimuli (Milner and Goodale 2006), pain (Melzack and Wall 1982) and threatening events (LeDoux 1996), blindsight (Weiskrantz 1990), covert facial recognition (Young and Burton 1999), and implicit memory (Schacter 1987). By extrapolating from such cases, one may speculate that nonconscious processing allows for a reasonable degree of functionality in 'zombie mode' outside artificially constructed experimental contexts. In turn, this raises the possibility that equivalent function performance may exist in the absence of consciousness. Chalmers takes this to entail that consciousness is not necessary for cognitive and behavioral performance, hence his claim that consciousness cannot be a problem about the performance of functions.

The possibility envisaged by Chalmers is, however, mitigated by the fact that the extrapolation on which it relies turns out to be false. Patients suffering from deficits in awareness invariably exhibit a significant degree of dysfunctionality under routine, everyday conditions:

One way to find out what something is good for is to examine what it is like not to have it. [...] there is a broad spectrum of syndromes in which there is a loss of acknowledged awareness of capacities or their contents, ranging from detection, through selective attention, semantic and associative meaning, episodic memory, to language. [...] The message that emerges from the clinic is unmistakable: all of the syndromes can possess implicit processing, but none of the patients can live by implicit processing alone. It cannot be used by the patient in thinking or in imagery, and this is a severe penalty. [...] The amnesic patient is severely impaired, and requires continuous custodial care. Priming is intact, but of no evident use to the amnesic victim. He cannot relate what is primed today to what was primed yesterday, or to any other item in memory, including time and place and other (but not only) contextual information; he is functionally fixed in the semantic or procedural present. [...] Similarly, the blindsight patient continues to fail to identify objects and to bump into them in his blind field. If he can detect a stimulus in the blind field, he does not know what it is. There may be some occasional benefit to him if he can duck as a rapidly zooming object approaches (although typically this is not a common response in blindsight subjects). The blindsight subject cannot image the stimulus, about which he has just guessed, in relation to other stimuli, or to their spatial setting, because it is not perceived." (Weiskrantz 1997, 168–69)

The recurring theme emerging from clinical observations is that patients behave as if covertly processed information is not processed at all. It is only from the external perspective of the experimenter that nonconscious processing can be evidenced and only under the external prompting of the experimenter that the patient may act on the information covertly processed. This phenomenon is particularly obvious in blindsight patients, who only exhibit good performance in the experimental context of forced-choice tasks prompting them to guess which option is correct. Left on their own devices, blindsight patients fail to spontaneously initiate visually guided behaviour in response to stimuli in their impaired visual field (Cowey 2010; Marcel 1983). This observation has been used to support the rival view that consciousness has a function in normal subjects. In particular, global workspace theories posit that "consciousness is required for some specific cognitive tasks, including those that require durable information maintenance, novel combinations

of operations, or the spontaneous generation of intentional behavior” (Dehaene and Naccache 2001, 1).

What does this entail for the functional definability of consciousness? So far, nothing conclusive. Global workspace theorists argue that consciousness has a function because loss of consciousness correlates with loss of task performance, while Chalmers argues that consciousness is not functionally definable because of observed and extrapolated dissociations between task performance and consciousness. In both cases, conclusions about functional definability are based on prior knowledge of association/dissociation, not of causation/absence of causation. The problem with such inferences is that association doesn’t always entail causation and causation doesn’t always entail association.

The global access hypothesis faces the difficulty of inferring causation from association. While there is ample empirical evidence demonstrating a robust association between loss of consciousness and loss of function, this is insufficient to demonstrate the causal link required for functional ascription. For instance, Zeki’s claim that the function of V5 is to detect motion is supported by experiments in which the item which is ascribed a function (i.e., V5) is known to be independently manipulated. In contrast, in the experiments cited in support of the conscious access hypothesis, it is not clear that consciousness is independently manipulated. The awareness deficits (amnesia, blindsight, prosopagnosia, aphasia, etc.) mentioned by Weiskrantz involve natural experiments in which what is in fact manipulated (the independent variable) is brain activity, not consciousness. Assuming that brain lesion patients are comparable to healthy subjects in all respects except for localized loss of brain activity, it can be inferred that the lesioning of specific brain areas causes both loss of awareness and loss of performance for certain types of tasks. However, nothing here justifies the additional inference that loss of consciousness is responsible for loss in performance (Block 1995). Strictly speaking, these natural experiments only show that brain lesions are causally relevant to both awareness and task performance. It could be that consciousness has the function global workspace theories attribute it (Baars 2002; Dehaene and Naccache 2001), just as it could be that functional correlates are required for stimulus awareness, as postulated by the feature integration theory (Treccani 2018; Treisman and Gelade 1980). Or again, according to an interpretation compatible with Chalmers’s antifunctionalism, consciousness and its functional correlates could be divergent effects of a common cause (LeDoux 1996; LeDoux and Pine 2016). Currently accepted standards of internal validity dictate that singling out the correct causal explanation of an observed correlation requires studies capable of ruling out rival interpretations; in turn, this requires an experiment in which consciousness is independently manipulated.²

Chalmers’s functional undefinability thesis faces the converse difficulty of inferring lack of causation given lack of association. Just as the knockout of a gene may result in no phenotypic differences because a second gene takes over the function of the knocked-out gene, it is conceivable that a factor or mechanism *Z* compensates for the loss of performance caused by the loss of consciousness. In this scenario, both consciousness and *Z* are causally relevant, and therefore play a functional role vis-à-vis task performance, but since the inhibitory effect of consciousness knockout

²In order to demonstrate that a variable of interest is causally relevant to a correlated outcome, a controlled experiment satisfying three desiderata is typically required (Baetu 2020; Shadish et al. 2002; Woodward 2003): (i) An intervention on the variable (i.e., an experiment) must be conducted. Manipulation—as opposed to mere observation of differences in the outcome between two conditions—is standardly required in order to establish the directionality of causation (i.e., demonstrate that the variable is causally relevant to the outcome rather than the other way around) and rule out the possibility that the changes in the variable and the outcome are correlated due to a common cause. (ii) The test and the control conditions must be comparable in all relevant respects except for the variable manipulated in the experiment. Failure to ensure comparability raises the possibility that some other difference between the two conditions (a confounder) is responsible for the observed differences in outcomes. (iii) The intervention should be accurate in the sense that it should target only the variable under investigation. Accuracy is required to demonstrate the causal relevance of the tested (independent) variable to the differences in outcomes. If the accuracy of the intervention cannot be demonstrated, the causal efficacy of the intervention may be attributed to the fact that the intervention targets some other variable in addition to or instead of the tested variable.

is masked by the excitatory effect of *Z*, no difference in task performance is observed when comparing performance in zombie and normal subjects. This indicates that in order to infer lack of causation/function, one cannot solely rely on actual or conceivable dissociations between consciousness and function; additional information about how data was generated needs to be known or assumed.³

2.d The functional undefinability criterion lacks specificity

To accurately discriminate between hard and easy problems, it is not enough to show that consciousness is not functionally definable (i.e., demonstrate criterion sensitivity); it must also be shown that all or most other phenomena are functionally definable (demonstrate specificity). Without such evidence, the criterion may correctly identify consciousness (the true positive) along with a lot of everything else (false positives) as hard problems.⁴ Is it true then that all/most other biological and psychological phenomena are functionally definable?

A causal role understanding of functions entails that the only things that are not functionally definable are those which are never causes—that is, epiphenomena. These are not as rare as Chalmers seems to assume. Many phenomena in biology and psychology amount to stimulus-response causal sequences replicated and studied in the laboratory (Baetu 2019; Bechtel and Richardson 2010; Craver 2007; Darden 2006). In most cases, when a phenomenon is replicated in the laboratory, it is cut off from the causal structure of the world since whatever causal role the response may play under natural conditions is not allowed to unfold; hence, for all intents and purposes, the response studied is functionally undefined. Moreover, some aspects of a response may be always epiphenomenal, and therefore functionally undefinable. For instance, sunburns are experimentally characterized as the stimulus-response phenomenon of ultraviolet radiation-induced erythema (a type of inflammatory response) (Rainsford 2015). The most distinctive feature of sunburns is the redness of the skin (erythema), which is used to measure the magnitude of the inflammatory response. Yet it is very likely that the redness itself doesn't play any causal role (physiological, evolutionary, or other); it is simply a sterile side effect of increased blood flow. Such counterexamples undermine the assumption that all or almost all biological and psychological phenomena are functionally defined.

2.e Functional undefinability is not an obstacle to mechanistic explanation

The second axis of the functional definability criterion is the link between functional definability and mechanistic explainability. According to Chalmers, it is a conceptual truism that all it could possibly take to explain the performance of a function, and thus solve an easy problem, is to specify a mechanism. This, however, seems doubtful. For instance, in the case of a power hammer, the variable 'mass of the hammer' is causally relevant to the outcome 'force applied on the target,' as determined by controlled experiments with different loads. The variable 'mass of the hammer' is therefore functionally definable, yet there is no mechanism linking mass and force. Perhaps Chalmers refers to the fact that most phenomena in the life sciences are characterized as functional 'black boxes' linking stimuli to responses whose inner workings are subsequently elucidated by specifying mechanisms (Baetu 2019; Bechtel and Richardson 2010; Craver and Bechtel 2007;

³In causal modelling, instances of silent causation are ruled out by the minimality assumption, which is meant to ensure that all causal relationships describing a target of interest (e.g., physical system, experiment) can be mathematically represented as probabilistic dependencies between variables (Pearl et al. 2016). In experimental research, evidence for lack of association is interpreted in light of background knowledge (Baetu 2022). For example, in clinical and public policy contexts, such evidence supports the pragmatically correct conclusion that the tested intervention doesn't make any causal difference to the outcome. In genetics, experiments showing that knocking out a gene doesn't result in phenotypic differences typically prompt a search for homologous DNA sequences.

⁴Kahneman (2011) discusses examples of fallacies resulting from a failure to take into account test specificity.

Machamer et al. 2000). If so, then this is not a conceptual truism, but a contingent, empirical fact about the life sciences.

Conversely, Chalmers holds that the set of functionally undefinable phenomena—which he takes to specifically include only or almost only consciousness—are not mechanistically explainable. This, too, is doubtful. Functionally undefinable phenomena as defined by Chalmers are ‘epiphenomena.’ But any effect, epiphenomenal or not, can be in principle explained by a causal mechanism. For example, even if erythema is most probably a side effect devoid of any functional relevance, it is mechanistically explained by a local increase of blood flow. Moreover, from an epistemological point of view, epiphenomenalism facilitates the elucidation of mechanisms. Ideally, what happens or doesn’t happen after a response is generated during the replication of stimulus-response phenomenon is not part of the mechanistic explanation of that phenomenon. This requirement is satisfied if nothing further happens, if the response never feeds back into the mechanism, or if the feedback occurs at a different timescale than that of the phenomenon of interest.⁵ What worries scientists is not that epiphenomenalism may be true. Quite on the contrary, their main concern is that what is treated as an epiphenomenal response in the laboratory may not be so under natural conditions and that this may have an impact on the physiological relevance of the proposed mechanism. Downstream effects may feed back into the mechanism, altering its structure and dynamics at physiologically relevant timescales. Such feedback loops turn any attempt to quantitatively model and predict the states and outcomes of the mechanism into a nightmare of partial differential equations (Shmulevich and Aitchison 2009). In contrast, what doesn’t happen after an epiphenomenon is generated is a state of perpetual nonhappening that doesn’t require any further explanation or modelling. Thus, that consciousness may not play a causal role, and thus fail to have a function in the context of a system doesn’t constitute an obstacle to mechanistic explanation; if anything, it should facilitate the explanatory project.

3. The criterion of the ‘accompanying phenomenon’

3.a *The subjective experience accompanying cognitive and behavioral functions*

The criterion of functional definability is meant to capture consciousness, functional undefinability and mechanistic unexplainability under the category ‘hard problem of consciousness,’ and all/most other biological and psychological phenomena, functional definability and mechanistic explainability under the category ‘easy problems.’ It turns out, however, that the criterion lacks the perfect sensitivity and near-perfect specificity Chalmers attributes it. The claim that consciousness is functionally undefinable is unjustified, and even if sensitivity were proven, functional undefinability would still lack the required specificity to accurately discriminate consciousness from epiphenomena such as erythema. The fact that functional definability doesn’t guarantee mechanistic explainability and functional undefinability doesn’t entail mechanistic unexplainability further undermines the accuracy of the criterion as a tool for discriminating between the hard problem of consciousness and the easy problems of [almost] everything else.

Notwithstanding, Chalmers has a fallback position:

This is not to say that experience has no function. Perhaps it will turn out to play an important cognitive role, but for any role it might play, there will be more to the explanation of experience than a simple explanation of the function. (2010, 8)

⁵The last scenario refers to a distinction between a proximal explanation of the ‘mechanistic how’ filling in the details of the causal processes linking stimulus and response, and the causal implications of the response in some wider context, most notably its impact on biological fitness, which is part of the ultimate explanation of the ‘evolutionary why’ (Mayr 1961).

What makes the hard problem hard and almost unique is that it goes *beyond* problems about the performance of functions. To see this, note that even when we have explained the performance of all the cognitive and behavioural functions in the vicinity of experience—perceptual discrimination, categorization, internal access, verbal report—a further unanswered question may remain: *Why is the performance of these functions accompanied by experience?* (2010, 8)

The allocution ‘in the vicinity of’ refers to a measure of statistical dependence, in this case the association between conscious experience and functional correlates:

The easy problems of consciousness include those of explaining the following phenomena: the ability to discriminate, categorize, and react to environmental stimuli; the integration of information by a cognitive system; the reportability of mental states; the ability of a system to access its own internal states; the focus of attention; the deliberate control of behaviour; the difference between wakefulness and sleep. All of these phenomena are associated with the notion of consciousness. (2010, 4)

Thus, consciousness “goes *beyond* problems about the performance of functions,” in the sense that it correlates with certain cognitive and behavioral functions, yet an explanation of these functions doesn’t explain consciousness.

Chalmers takes this ‘association/accompanying without explanation’ criterion to single out a highly distinctive feature of consciousness, and argues that a conceptual mistake test demonstrates the accuracy of the criterion:

If someone says, “I can see that you have explained how DNA stores and transmits hereditary information from one generation to the next, but you have not explained how it is a gene,” then they are making a conceptual mistake. All it means to be a gene is to be an entity that performs the relevant storage and transmission function. But if someone says, “I can see that you have explained how information is discriminated, integrated, and reported, but you have not explained how it is experienced” they are not making a conceptual mistake. (2010, 8)

3.b The sensitivity of the criterion of the accompanying phenomenon is uncertain

Proof of sensitivity rests entirely on the claim that no conceptual mistake is made when one says that consciousness remains unexplained even if functional correlates of consciousness are explained. Presumably, Chalmers takes this to follow from the tacit assumption that it is false that all it means to be conscious is to perform the functions of discrimination, integration, and reporting of information. But what justifies this assumption? Two possible responses may be envisaged. The first would be to treat this assumption as a definitional matter. If we go down this path, we run into the problem of conceptual relativism. For example, since Tononi and Koch (2015) are committed to the view that consciousness is nothing else but integrated information, they would insist that one does make a conceptual mistake when claiming that an explanation of how information is discriminated and integrated leaves consciousness unexplained.

A more promising option is to seek an empirical justification all parties might accept. For example, LeDoux (1996) hypothesized that adaptive behavioral responses correlate with conscious feelings of fear because the two are divergent effects of a common cause, namely threatening stimuli. A common cause model entails that an explanation of the mechanism linking stimulus and behavioral correlate will not shed any light on the mechanism linking stimulus and conscious experience. If this model turns out to be correct and can be generalized to other functional correlates, then one can conclude that, given our current understanding of causation, no conceptual mistake is made by stating that an explanation of functional correlates will not explain

consciousness. For the time being, however, there is no conclusive evidence to support this view. As discussed in [section 2.c](#), currently available evidence is likewise compatible with rival models hypothesizing that functional correlates causally determine consciousness—as, for instance, proposed by the feature integration theories (Trecani 2018). According to these models, the mechanisms underpinning functional correlates overlap with those of consciousness. If we accept these models, it would be a conceptual mistake to expect that an explanation of functional correlates will contribute nothing to an explanation of consciousness or to rule out the possibility that an explanation of functional correlates may suffice to explain consciousness.

Whether we pursue the definitional route or the empirical alternative, the conceptual mistake test is inconclusive: consciousness may or may not be said to “go beyond” problems about the performance of functions depending on background assumptions about how consciousness is defined and which causal model best explains the association between consciousness and functional correlates.

3.c The criterion of the accompanying phenomenon lacks any usable degree of specificity

Before proceeding to an evaluation of the specificity of the conceptual mistake test, some preliminary clarifications are needed. First, the gene is neither a phenomenon in need of an explanation, nor an explanation of anything, but a scientific concept defining a class of objects. A concept refers or fails to refer depending on whether there are things in the world that correspond to the description postulated by the concept, and the classifications it generates are objective or subjective depending on whether the things to which it refers constitute a natural kind (Boyd 2010; Machery 2009). Second, when Chalmers states that “all it means to be a gene is to be an entity that performs the relevant storage and transmission function,” he is not talking about genes as understood in genetics and molecular biology, but of a functionalist gene* concept of his own invention.⁶ Third, Chalmers’s gene* concept refers indiscriminately to chromosomes, plasmids, maternal RNA, transcription factors and their cellular localization, DNA methylation and histone acetylation, mitochondria and mitochondrial DNA, as well as a host of environmental influences (drugs, pathogens, viruses, prions) transmitted vertically or horizontally. Since a different mechanism underpins a different pattern of heredity in each case, this collection of factors and processes doesn’t constitute a natural kind.

This last point indicates that even if s_1 and s_2 are genes* (i.e., storers and transmitters of inheritance information), explaining how s_1 is a gene* may not shed any light on how s_2 is a gene*. For instance, one doesn’t make a conceptual mistake in saying, “I can see that you have explained how chromosomal DNA is a gene*, but you have not explained how maternal RNA, is a gene*.” It should likewise be clear that the fact that an item s_i falls within the extension of a concept C_1 doesn’t necessarily entail that s_i doesn’t also fall within the extension of some other concepts $C_2, C_3, \dots C_j$. Thus, one can say, “I can see that you have explained how s_i is a C_1 , but you have not explained how s_i is a $C_2, C_3, \dots C_j$.” This is particularly obvious in the case of functional concepts: since any given cause can have many effects, it can play multiple causal roles, and thus have multiple functions, each captured by a distinct functional concept. For instance, since DNA performs the relevant storage and transmission function, DNA is a gene*. However, DNA also plays a role in maintaining the structural integrity of chromosomes. Thus, one doesn’t make a conceptual mistake in saying, “I can see that you have explained how DNA stores and transmits hereditary information from one

⁶In classical genetics, a gene is a chromosomal locus associated with a difference in phenotype, as determined by a genetic mapping technique. In molecular biology and genomics, a gene is an open reading frame, a transcription unit, or, more generally, a set of instructionlike DNA sequences known or expected to interact with the genome expression machinery of the cell. These two gene concepts are associated with explanations of two distinct aspects of heredity, namely patterns of inheritance and genotype-phenotype correlations, and refer to overlapping, but not identical, sets of objects (Baetu 2010; Griffiths and Stotz 2013).

generation to the next, but you have not explained how it is a structure supporting the chromosome.” Finally, a functionalist concept of the form ‘All it means to be X is to perform function F ’ fixes only the functional outcome without positing any constraints on the mechanistic how causing the outcome. In particular, such concepts don’t prohibit a priori mechanisms involving conscious performers of a function. Thus, the protagonist of a dystopian novel wouldn’t commit a conceptual mistake by saying, “I can see that you have explained how downsized nanoproletarians store and transmit hereditary information from one generation to the next, but you have not explained how they experience their condition.”

Whichever way we turn it, the conceptual mistake test fails to deliver any kind of specificity. No conceptual mistake is made if one claims that the explanation of how a gene* is a gene* doesn’t explain how an[other] gene* is a gene*, or how a gene* does something else than perform the function of gene*, including how a gene* is conscious. In response, some may be tempted to retort that the conceptual mistake test might fare better for other concepts or functions. I have serious reserves about this rebuttal. For one thing, the objections I raised generalize to any concept that categorizes things based on functional attribution alone. A deeper concern is that Chalmers commits a category mistake when he proposes to compare a phenomenon to a concept, be it functional or not. Since a hard problem is a mechanistically unexplainable phenomenon and an easy problem is a mechanistically explainable phenomenon, a test should assess whether the criterion of the accompanying phenomenon discriminates between mechanistically explained phenomena and the phenomenon of consciousness, and not whether it discriminates between concepts and the phenomenon of consciousness.

If we correct this inconsistency and consider the relevant term of comparison, it can be easily shown that many phenomena go beyond explanations of a given set of functions. For example, in the 1960s, immunologists observed that the phenomenon of sunburns is accompanied by several biological activities (Rainsford 2015). Yet even though the science of that time explained the performance of biological activities in the vicinity of sunburns, such as gene expression regulation and protein synthesis, a further unanswered question remained: Why was the performance of these activities accompanied by sunburns? In fact, there is an endless list of such ‘why’ questions. Why is λ -phage replication accompanied by cosmic background radiation? Why are religious beliefs accompanied by λ -phage replication? Why are lamps switches accompanied by religious beliefs? There is nothing absurd or contradictory in asking why statistically significant associations occur. Quite on the contrary, since such associations cannot be attributed to chance alone, they are paradigmatic examples of phenomena demanding an explanation (Baetu 2019; Bogen and Woodward 1988).

It can further be shown that at least some phenomena that go beyond explanations of a given set of functions are amenable to mechanistic explanations. In the case of sunburns, the mechanisms underpinning inflammatory responses were subsequently elucidated, revealing that the biological activities of gene expression regulation and protein synthesis, along with other mechanistic details, are causally involved in the production of sunburns (Clydesdale et al. 2001). This addresses the ‘why’ question raised earlier: sunburns accompany biological activities because the latter are causally relevant to the former. Yet even though the question has been answered and the mechanisms responsible for the phenomenon of sunburns have been explained, it is still the case that questions about sunburns go beyond problems about the performance of the functions of gene expression regulation and protein synthesis. This is simply because the latter are part of the mechanism of sunburns but are insufficient to explain the phenomenon. As for the other, seemingly more outlandish, ‘why’ questions mentioned above, they are banal cases of phenomena accompanying one another in virtue of having a common cause. Thus, the mere fact that a phenomenon goes beyond the performance of functions in their vicinity is neither a unique property of consciousness, nor a reliable indicator that researchers are confronted with a mechanistically intractable hard problem.

4. The ‘first-person data’ criterion

4.a *Objective functions vs. subjective data*

The last criterion of demarcation between hard and easy problems hinges on the distinction between the subjective nature of first-person data vs. the objective nature of third-person data:

to explain third-person data, we need to explain the objective functioning of a system and can do so in principle by specifying a mechanism. When it comes to first-person data, however, this model breaks down. The reason is that first-person data—the data of subjective experience—are not data about objective functioning. Merely explaining the objective functions does not explain subjective experience. (Chalmers 2010, 39)

The above proposes an alignment of third-person data with ‘being about objective functioning’/mechanistic explainability, and of first-person data with ‘being about subjective experience’ and ‘not being about objective functioning’/mechanistic unexplainability. We can safely conjecture that ‘objective functioning’ refers to the causal link between the thing that is attributed a function (the cause) and its function (to bring about an effect). This reading is consistent Chalmers’s definition of ‘function’ as a causal role (section 2.b), as well as with the assumption that, at least in the life sciences, functions are typically explained by mechanisms (section 2.e). As for the concept of ‘data being about something,’ it seems reasonable to assume that data are informative of (e.g., measure, predict, correlate with) the something they are about. Under this interpretation, first-person data are characterized as informative of subjective experience and uninformative of mechanistically explainable functional relationships, while third-person data are thought to convey information about mechanistically explainable functional relationships. As discussed in section 2, Chalmers takes functional definability and mechanistic explainability to be characteristic of the easy problems, while the absence of these properties is the distinctive mark of a hard problem.

4.b *The hard problem of consciousness is not uniformly hard*

A noticeable peculiarity of the above-proposed alignment is that while first-person data are about subjective experience and not about objective functioning, it is only stipulated that third-person data are about objective functioning; nothing is said about third-person data not being about subjective experience. Presumably, this omission is meant to accommodate the widespread assumption that whenever we talk about what we see, feel, think, and so on, we engage in spontaneous report experiments generating third-person data meant to inform—and occasionally misinform—others about the private first-person data of subjective experience. This assumption is empirically supported by self-experimentation studies in which researchers adopt simultaneously a first- and a third-person perspective (Head 1920; Piccinini 2009; Price and Aydede 2005; Price and Barrell 2012). Among other things, such studies provided evidence validating the use of verbal reporting as a method for assessing both consciousness and consciousness contents. In turn, verbal reports are commonly used to validate other measurement techniques, such as those relying on behavioral (Teasdale and Murray 2000), neurological (Owen et al. 2006) and informational (Casarotto et al. 2016) correlates of consciousness.

Researchers also attempted to disentangle the effects of conscious and nonconscious processing in verbal and verbal-like reporting. Experiments in which the reporting condition is manipulated revealed that the same stimulation condition may result in different discriminability thresholds⁷,

⁷Contemporary psychophysics relies on signal detection theory (Hautus et al. 2021), which draws a distinction between stimuli, internal representations (conscious or not) of stimuli, and decision biases. It is assumed that variability in the exposure procedure and the nervous system makes it such that a range of sensory values are elicited from one instance of the same stimulus to another. Thus, a subject’s ability to discriminate between two stimuli (denoted by the model parameter ‘sensitivity,’ or d') is dependent on both how different the two stimuli are and how accurately these differences are represented internally by

depending on the instructions used to elicit reports (e.g., report, guess, rate confidence, introspect, etc.), the modality of the report (verbal, pointing, clicking, blinking) and other constraints imposed on reporting (e.g., leisurely vs. as fast as possible; yes/no forced choice vs. yes/no/I don't know). By far, the best understood dissociation concerns 'guess' vs. 'awareness' verbal reports (Dienes 2008; Dienes et al. 1995), the contrasting of which emerged as the method of choice for demonstrating "nonconscious perception in both normal subjects and blindsight patients" (Marcel 1994, 80). A statistically significant distinction can be made between a 'subjective' and an 'objective discriminability threshold' corresponding to "detection level at which subjects claim not to be able to discriminate perceptual information at better than a chance level" [when instructed to report awareness] vs. "detection level at which perceptual information is actually discriminated at a chance level" [when subjects are instructed to guess] (Cheesman and Merikle 1984, 391). According to a popular interpretation, the subjective threshold is neither about the stimulus, nor about the subject's perceptual acuity and its impact on task performance, which are measured by the objective threshold, but rather about the subject's conscious awareness of the stimulus and decisions about how to act on this awareness.

If we accept adequately controlled verbal reports and other correlates as legitimate examples of third-person data about consciousness, then, according to the first- vs. third-person data demarcation criterion, some third-person data are both about consciousness and about mechanistically explainable objective functions. But if this is the case, then that at least some aspects of consciousness—namely those measured by third-person data—can be explained mechanistically whether or not first-person data are amenable to mechanistic explanation.

4.c The partial transparency of experience

In response to this caveat, Chalmers may turn tables and insist that whether or not third-person data about consciousness are available and explainable, first-person data of subjective experience 'go beyond' problems about objective functioning, that these data remain unexplained, and that the hard problem of consciousness is precisely about the unexplained character of first-person data. Yet this cannot possibly be true of all first-person data. Being informative about something is a symmetrical relationship: if *A* is informative of *B*, *B* is necessarily informative of *A* (Steinhart 2018, ch. 6). Thus, if in addition to being informative of objective functioning, some third-person data also convey information about the first-person data of subjective experience, then some first-person data must likewise convey information about things measured by third-person data, namely objective experience, in addition to being about subjective experience. In such cases, the properties 'being about subjective experience' and 'being about objective functioning' cannot be exclusively assigned to first- and third-person data, respectively.

It is not difficult to empirically verify that this is indeed the case. At least some first-person subjective experiences, such as feeling of color and motion, are informative of intersubjectively verifiable causal phenomena such as changes in color of a litmus paper and the motion of the moon in the sky.⁸ But if first-person data are not exclusively aligned with subjective experience/the

the nervous system. *d'* reflects the subject's true perceptual acuity, that is, the information made available to the subject upon exposure to the stimulus and based on which the subject can make a decision (e.g., to report whether a given stimulus was perceived or not). In contrast, decision bias ('criterion,' or *c*) tells researchers how the subject chooses to use this information. Given assumptions or prior knowledge about the probability distributions associated with the subject's internal representations, it is possible to calculate the model parameters *d'* and *c* directly from empirical data on hit and false alarm rates.

⁸In philosophy of science, subjective experiences informative of the external world are known as 'direct observations' and are construed as instrument-unaided measurements (Carnap 1928; van Fraassen 1980). In philosophy of mind, such experiences are taken to justify the 'transparency of experience' thesis (Tye 2002). Note, however, that although subjective experiences are routinely used in science as a means of measuring the world, figuring out what properties of the world qualia actually track is by no means obvious; see, for instance, the case of perceived colour (Zeki 1993).

mechanistically intractable hard problem of consciousness and third-person data are not exclusively aligned with objective functioning/mechanistically tractable easy problems, then the criterion is noncategorical: at least for some first- and third-person data, it classifies problems as simultaneously hard and easy.

Moreover, it is impossible to determine a priori whether any given first-person datum is solely about subjective experience or about subjective experience and objective functioning. Here is an example that illustrates the problem. Pain and erythema are assessed by essentially the same measurement technique. In clinical practice, pain is usually measured by instructing patients to introspect and rate their current pain levels on a numerical scale of 0 (no pain) to 10 (worst pain imaginable) (Noble et al. 2005). In a similar way, clinicians usually measure the intensity of skin inflammation by visual assessment and reporting a value on a four-point scale ranging from ‘no erythema’ to ‘violet erythema with edema’ (Rainsford 2015). From the first-person perspective of the subject assessing pain or erythema, measurements are based entirely on subjective experiences of what pain or visual awareness feel like now and how these subjective experiences compare with past and imagined experiences. From this perspective, there are no reasons to suspect that erythema measurements are any more or less likely to be about objective functioning than pain measurements. It is only after measurements are reported and shared intersubjectively that the subject finds out that some of the reported experiences systematically agree with the reports issued by other agents, while others don’t. In this particular example, an intersubjective agreement is systematically reached about erythema ratings—as indicated by the reliability of the scale, as well as consistency with instrument-based measures (Fullerton et al. 1996). In contrast, it is much more common that pain ratings are not amenable to intersubjective agreement or consistent with behavioral and physiological measures, suggesting that pain assessment can generate information of a strictly subjective nature (IASP Task Force on Taxonomy 2020).⁹

If it cannot be established a priori to what extent first-person experience is amenable to intersubjective agreement, then the premise that “that first-person data are not data about objective functioning” is not a self-evident or necessary truth, but an empirical fact contingent on what happens after any given first-person experience is shared. Since this premise is taken to justify the conclusion that “explaining the objective functions does not explain subjective experience,” it can further be objected that this conclusion may or may not be justified depending on whether the justificatory premise is true or not.¹⁰ In other words, the overall ‘easiness’ or ‘hardness’ of the problem of consciousness is contingent on the extent to which first- and third-person data are shown to convey or fail to convey information about one another.

⁹This potential lack of intersubjective agreement is epitomized by McCaffery’s operationalized definition, “Pain is whatever the experiencing person says it is, existing whenever he says it does” (1968, 95), which is the current gold standard in clinical practice. Nevertheless, in some respects, even pain reports are amenable to a considerable degree of intersubjective agreement. For instance, when induced to elicit pain reports, the higher or lower values of the pain ratings can be predicted when subjects are presented with certain stimuli (e.g., noxious vs. non-noxious stimulation), information (e.g., raise or lower pain expectation), in the context of certain behaviors (e.g., when distracted), and when administered analgesics (Hewer et al. 1949). Moreover, certain forms of noxious pain are reliably predicted by patterns of brain activity (Wager et al. 2013). Statistical analyses of descriptive terms appearing in verbal reports of pain revealed a reproducible structure of the space of pain experiences, such as dimensionality (ways in which a pain can vary) (Melzack and Casey 1968) and correlations between pain descriptors (Melzack 1975).

¹⁰Presumably, the justification of this claim has more to do with a lack of understanding of how first-person data are generated—notably, how introspection works as an ‘instrument of measurement’ of mental states (Marcel 1993; Overgaard and Sørensen 2004)—than to what data are about. However, this problem is not exclusive to consciousness research. In general, the explanation of a phenomenon is not the same thing as and does not entail an explanation of how the methods used to measure and probe that phenomenon work. For instance, an explanation of ultraviolet radiation-induced erythema (sunburns) doesn’t explain how erythema is assessed by experiencing subjective appearances. The same is true of the explanation of DNA replication in terms of a semiconservative mechanism proposed by Watson and Crick (1953), which doesn’t explain how Meselson and Stahl (1958) measured the relative ratios of parent and daughter DNA strands.

5. Conclusion

Chalmers argues that while all/almost all phenomena in the life sciences are easy problems that can be mechanistically explained, the phenomenon of consciousness constitutes a mechanistically intractable hard problem. This claim presupposes, first, that a principled, criterion-based distinction can be made between easy and hard problems and, second, that once the criteria are applied to various phenomena, consciousness systematically falls in the class of hard problems while all/almost all other phenomena fall into the class of easy problems. The evaluation conducted in this paper shows that none of the proposed criteria of demarcation between hard and easy problems succeeds in singling out consciousness as a unique, mechanistically unexplainable phenomenon. I conclude therefore that Chalmers fails to identify a unique property of the phenomenon of consciousness that may allow us to infer, prior to any further scientific investigation, that consciousness will forever hover as an unexplainable phenomenological surplus over and above a mechanistic understanding of living organisms.

Acknowledgements. This research was supported by SSHRC Grant # 430-2020-0654.

Tudor Baetu is associate professor of philosophy of science at the Université du Québec à Trois-Rivières. His research interests include the epistemology and metaphysics of causal-mechanistic explanations, the explanatory role of mathematical models and computer simulations, and methodological issues in experimental science.

References

- Baars, Bernard J. 2002. "The Conscious Access Hypothesis: Origins and Recent Evidence." *Trends in Cognitive Sciences* 6 (1): 47–52.
- Baetu, Tudor M. 2010. "The Referential Convergence of Gene Concepts Based on Classical and Molecular Analyses." *International Studies in the Philosophy of Science* 24 (4): 411–27.
- Baetu, Tudor M. 2019. *Mechanisms in Molecular Biology*. Part of the *Elements in the Philosophy of Biology* series, edited by Grant Ramsey and Michael Ruse. Cambridge: Cambridge University Press.
- Baetu, Tudor M. 2020. "Causal Inference in Biomedical Research." *Biology and Philosophy* 35: 43.
- Baetu, Tudor M. 2022. "Inferential Pluralism in Causal Reasoning from Randomized Experiments." *Acta Biotheoretica* 70: 22.
- Bechtel, William, and Adele Abrahamsen. 2005. "Explanation: A Mechanist Alternative." *Studies in History and Philosophy of Biological and Biomedical Sciences* 36: 421–41.
- Bechtel, William, and Robert Richardson. 2010. *Discovering Complexity: Decomposition and Localization as Strategies in Scientific Research*. Cambridge, MA: MIT Press.
- Block, Ned 1995. "On a Confusion about a Function of Consciousness." *Behavioral and Brain Sciences* 18: 227–47.
- Bogen, James, and James Woodward. 1988. "Saving the Phenomena." *The Philosophical Review* 97 (3): 303–52.
- Boyd, Richard. 2010. "Realism, Natural Kinds, and Philosophical Methods." In *The Semantics and Metaphysics of Natural Kinds*, edited by Helen Beebe and Nigel Sabbarton-Leary, 212–34. New York: Routledge.
- Carnap, Rudolf. 1928. *The Logical Structure of the World*. Berkeley: University of California Press.
- Casarotto, Silvia, Angela Comanducci, Mario Rosanova, Simone Sarasso, Matteo Fecchio, Martino Napolitani, Andrea Pigorini, Adenauer G. Casali, Pietro D. Trimarchi, Melanie Boly, Olivia Gosseries, Olivier Bodart, Francesco Curtò, Cristina Landi, Maurizio Mariotti, Guya Devalle, Steven Laureys, Giulio Tononi, and Marcello Massimini. 2016. "Stratification of Unresponsive Patients by an Independently Validated Index of Brain Complexity." *Annals of Neurology* 80 (5): 718–29.
- Chalmers, David. 1996. *The Conscious Mind: In Search of a Fundamental Theory*. Oxford: Oxford University Press.
- Chalmers, David. 2010. *The Character of Consciousness*. Oxford: Oxford University Press.
- Cheesman, Jim, and Philip M. Merikle. 1984. "Priming with and without Awareness." *Perception and Psychophysics* 36 (4): 387–95.
- Clydesdale, Gavin J., Geoffrey W. Dandie, and H. Konrad Muller. 2001. "Ultraviolet Light Induced Injury: Immunological and Inflammatory Effects." *Immunology and Cell Biology* 79: 547–68.
- Cowey, Alan. 2010. "The Blindsight Saga." *Experimental Brain Research* 200 (1): 3–24.
- Craver, Carl. 2001. "Role Functions, Mechanisms, and Hierarchy." *Philosophy of Science* 68: 53–74.
- Craver, Carl. 2007. *Explaining the Brain: Mechanisms and the Mosaic Unity of Neuroscience*. Oxford: Clarendon Press.
- Craver, Carl, and William Bechtel. 2007. "Top-Down Causation without Top-Down Causes." *Biology and Philosophy* 22: 547–63.
- Cummins, Robert. 1975. "Functional Analysis." *Journal of Philosophy* 72 (20): 741–65.
- Darden, Lindley. 2006. *Reasoning in Biological Discoveries: Essays on Mechanisms, Interfield Relations, and Anomaly Resolution*. Cambridge: Cambridge University Press.

- Dehaene, Stanislas, and Lionel Naccache. 2001. "Towards a Cognitive Neuroscience of Consciousness: Basic Evidence and a Workspace Framework." *Cognition* 79: 1–37.
- Dienes, Zoltán. 2008. "Subjective Measures of Unconscious Knowledge." *Progress in Brain Research* 168: 49–64.
- Dienes, Zoltán, Gerry T. M. Altmann, Liam Kwan, and Alastair Goode. 1995. "Unconscious Knowledge of Artificial Grammars Is Applied Strategically." *Journal of Experimental Psychology: Learning, Memory, & Cognition* 21: 1322–38.
- Fullerton, A., T. Fischer, A. Lahti, K.-P. Wilhelm, H. Takiwaki, J. Serup. 1996. "Guidelines for Measurement of Skin Colour and Erythema. A Report from the Standardization Group of the European Society of Contact Dermatitis." *Contact Dermatitis* 35 (1): 1–10.
- Glennan, Stuart. 2017. *The New Mechanical Philosophy*. New York: Oxford University Press.
- Griffiths, Paul, and Karola Stotz. 2013. *Genetics and Philosophy: An Introduction*. Cambridge: Cambridge University Press.
- Hautus, Michael J., Neil A. Macmillan, and C. Douglas Creelman. 2021. *Detection Theory: A Users Guide*. New York: Routledge.
- Head, Henry. 1920. *Studies in Neurology*. London: Oxford University Press.
- Hewer, A. J. H., C. A. Keele, K. D. Keele, and P. W. Nathan. 1949. "A Clinical Method of Assessing Analgesics." *Lancet*: 431–35.
- IASP Task Force on Taxonomy *Pain Terms and Definitions* 2020 [cited]. <https://www.iasp-pain.org/resources/terminology>.
- Kahneman, Daniel. 2011. *Thinking, Fast and Slow*. New York: Farrar, Straus and Giroux.
- LeDoux, Joseph E. 1996. *The Emotional Brain: The Mysterious Underpinnings of Emotional Life*. New York: Simon & Schuster.
- LeDoux, Joseph E., and D. S. Pine. 2016. "Using Neuroscience to Help Understand Fear and Anxiety: A Two-System Framework." *American Journal of Psychiatry* 173 (11): 1083–93.
- Machamer, Peter, Lindley Darden, and Carl F. Craver. 2000. "Thinking about Mechanisms." *Philosophy of Science* 67 (1): 1–25.
- Machery, Edouard. 2009. *Doing without Concepts*. New York: Oxford University Press.
- Marcel, Anthony J. 1983. "Conscious and Unconscious Perception: An Approach to the Relations between Phenomenal Experience and Perceptual Processes." *Cognitive Psychology* 15: 238–300.
- Marcel, Anthony J. 1993. "Slippage in the Unity of Consciousness." In *Experimental and Theoretical Studies of Consciousness*, edited by Ciba Foundation Symposium 174. Chichester: Wiley.
- Marcel, Anthony J. 1994. "What Is Relevant to the Unity of Consciousness?" *Proceedings of the British Academy* 83: 79–88.
- Mayr, Ernst. 1961. "Cause and Effect in Biology." *Science* 134 (3489): 1501–6.
- McCaffery, Margo. 1968. *Nursing Practice Theories Related to Cognition, Bodily Pain, and Man-Environment Interactions*. Los Angeles: Regents of the University of California.
- McLaughlin, Brian. 2006. "Is Role-Functionalism Committed to Epiphenomenalism?" *Consciousness Studies* 13 (1–2): 39–66.
- Melzack, Ronald. 1975. "The McGill Pain Questionnaire: Major Properties and Scoring Methods." *Pain* 1: 277–99.
- Melzack, Ronald, and Kenneth L. Casey. 1968. "Sensory, Motivational, and Central Control Determinants of Pain: A New Conceptual Model." In *The Skin Senses*, edited by D. Kenshalo, 423–43. Springfield, IL: Thomas.
- Melzack, Ronald, and Patrick D. Wall. 1982. *The Challenge of Pain*. London: Penguin Books.
- Meselson, Matthew, and Franklin W. Stahl. 1958. "The Replication of DNA in *Escherichia coli*." *Proceedings of the National Academy of Science* 44: 671–82.
- Millikan, Ruth. 1989. "In Defense of Proper Functions." *Philosophy of Science* 56 (2): 288–302.
- Milner, David, and Melvyn Goodale. 2006. *The Visual Brain in Action*. Oxford: Oxford University Press.
- Neander, Karen. 1991. "Functions as Selected Effects: The Conceptual Analyst's Defense." *Philosophy of Science* 58 (2): 168–84.
- Noble, Bill, David Clark, Marcia Meldrum, Henk ten Have, Jane Seymour, Michelle Winslow, and Silvia Paz. 2005. "The Measurement of Pain, 1945–2000." *Journal of Pain and Symptom Management* 29 (1): 14–21.
- Overgaard, Morten, and Thomas Sørensen. 2004. "Introspection Distinct from First Order Experiences." *Journal of Consciousness Studies* 11: 77–95.
- Owen, Adrian M., Martin R. Coleman, Melanie Boly, Matthew H. Davis, Steven Laureys, and John D. Pickard. 2006. "Detecting Awareness in the Vegetative State." *Science* 313 (5792): 1402.
- Pearl, Judea, Madelyn Glymour, and Nicholas Jewell. 2016. *Causal Inference in Statistics: A Primer*. Chichester: Wiley & Sons.
- Piccinini, Gualtiero. 2009. "First-Person Data, Publicity & Self-Measurement." *Philosophers' Imprint* 9 (9): 1–16.
- Price, Donald D., and Murat Aydede. 2005. "The Experimental Use of Introspection in the Scientific Study of Pain and Its Integration with Third-Person Methodologies: The Experiential-Phenomenological Approach." In *Pain: New Essays on Its Nature and the Methodology of Its Study*, edited by Murat Aydede, 243–73. Cambridge, MA: MIT Press.
- Price, Donald D., and James J. Barrell. 2012. *Inner Experience and Neuroscience: Merging Both Perspectives*. Cambridge, MA: MIT Press.
- Rainsford, Kim D. 2015. "History and Development of Ibuprofen." In *Ibuprofen: Discovery, Development and Therapeutics*, edited by Kim D. Rainsford, 1–21. Chichester: John Wiley & Sons.
- Schacter, Daniel L. 1987. "Implicit Memory: History and Current Status." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 13 (3): 501–18.
- Shadish, William R., Thomas D. Cook, and Donald T. Campbell. 2002. *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Boston: Houghton Mifflin.
- Sharma, R. P. 1973. "Wingless a New Mutant in *Drosophila melanogaster*." *Drosophila information service* 50: 134.

- Shmulevich, Ilya, and John Aitchison. 2009. "Deterministic and Stochastic Models of Genetic Regulatory Networks." *Methods in Enzymology* 467: 335–56.
- Steinhart, Eric. 2018. *More Precisely: The Math You Need to Do Philosophy*. Peterborough, Ontario: Broadview Press.
- Teasdale, G. M., and L. Murray. 2000. "Revisiting the Glasgow Coma Scale and Coma Score." *Intensive Care Medicine* 26 (2): 153–54.
- Tononi, Giulio, and Christof Koch. 2015. "Consciousness: Here, There and Everywhere?" *Philosophical Transactions of the Royal Society B* 370: 20140167.
- Treccani, Barbara. 2018. "The Neuropsychology of Feature Binding and Conscious Perception." *Frontiers in Psychology* 9 (2606): 1–5.
- Treisman, Anne, and Garry Gelade. 1980. "A Feature-Integration Theory of Attention." *Cognitive Psychology* 12 (1): 97–136.
- Tye, Michael. 2002. "Representationalism and the Transparency of Experience." *Noûs* 36 (1): 137–51.
- van Fraassen, Bas C. 1980. *The Scientific Image*. New York: Oxford University Press.
- Wager, Tor D., Lauren Y. Atlas, Martin A. Lindquist, Mathieu Roy, Choong-Wan Woo, and Ethan Kross. 2013. "An fMRI-Based Neurologic Signature of Physical Pain." *New England Journal of Medicine* 368 (15): 1388–97.
- Watson, James D., and Francis H. Crick. 1953. "Genetical Implications of the Structure of Deoxyribonucleic Acid." *Nature* 171 (4361): 964–67.
- Weiskrantz, Lawrence. 1990. *Blindsight: A Case Study and Implications*. New York: Oxford University Press.
- Weiskrantz, Lawrence. 1997. *Consciousness Lost and Found: A Neuropsychological Exploration*. Oxford: Oxford University Press.
- Woodward, James. 2003. *Making Things Happen: A Theory of Causal Explanation*. Oxford: Oxford University Press.
- Young, Andrew W., and A. Mike Burton. 1999. "Simulating Face Recognition: Implication for Modelling Cognition." *Cognitive Neuropsychology* 16: 1–48.
- Zeki, Semir. 1993. *A Vision of the Brain*. Oxford: Blackwell.
- Zeki, Semir. 2015. "Area V5: A Microcosm of the Visual Brain." *Frontiers in Integrative Neuroscience* 9 (21). <https://doi.org/10.3389/fnint.2015.00021>.