PRINCIPIA **29**(1): 39–72 (2025) doi: 10.5007/1808-1711.2025.e94823
Published by NEL — Epistemology and Logic Research Group, Federal University of Santa Catarina (UFSC), Brazil.

# WHY WE MUST CARE ABOUT ANIMAL CONSCIOUSNESS: AGAINST CARRUTHERS' NIHILISM

VICTOR MACHADO BARCELLOS
*Universidade Federal do Rio de Janeiro (UFRJ),* BRAZIL
victorbarcellos1995@gmail.com
https://orcid.org/0000-0003-2269-4923

**Abstract.** One of the most challenging positions in contemporary philosophy of animal consciousness is that proposed by Peter Carruthers (2018a, 2018b, 2019, 2020). According to Carruthers, there is no fact of the matter about whether animals instantiate conscious states. This radical conclusion arises from the conjunction of two theses he endorses: the global workspace theory and the phenomenal concept strategy. This paper argues against Carruthers' radical viewpoint. Its structure is as follows. First, I will present Carruthers' theses on consciousness, such as the all-or-nothing characterization of consciousness and the distinction between 'qualia realism' and 'qualia irrealism'. Subsequently, I will provide a brief overview of the global workspace theory and the phenomenal concept strategy. Next, I will reconstruct the arguments that underpin Carruthers' skepticism about attributing consciousness to animals. Finally, I will present two arguments that challenge Carruthers' position, highlighting inherent contradictions within his project. Contrary to Carruthers' controversial assertion (2020, p.18), I will conclude that animal consciousness deserves attention from both philosophy and the cognitive sciences.

**Keywords:** animal consciousness • science of consciousness • philosophy of consciousness • animal cognition • Carruthers • representation

RECEIVED: 07/06/2023 REVISED: 15/05/2024 ACCEPTED: 31/01/2025

## 1. Introduction

We are currently living in a golden age of research on phenomenal consciousness in non-human animals[1]. Specialized journals on the subject were created, and laws on animal welfare based on the instantiation of consciousness have been passed in several parliaments around the world (Browning & Veit 2022). The presence of consciousness in adult humans is largely evident, as we can readily attest to the unpleasantness of pain, the joy of listening to Mozart's *Die Zauberflöte*, and the distinct sensory experience of perceiving an apple. However, the question of whether other species share conscious states similar to ours remains less clear.[2]

Determining which animals are conscious is by no means a trivial task. It necessitates the development of a reliable and precise method capable of discerning its

presence in different populations of animals. How can we ascertain whether animals as distinct and distant on the phylogenetic tree of life as dolphins and hermit crabs possess consciousness? This issue is known as '*the measurement problem of animal consciousness*' (Browning & Veit 2020).[3] The measurement problem of animal consciousness involves developing an appropriate methodology for determining whether a given animal possesses consciousness. This issue is closely tied to the distribution question, which examines which organisms in the animal kingdom are conscious (Allen & Trestman 2016). However, to address the distribution question, a reliable method is required to determine whether a given animal species instantiates conscious mental states. Consequently, to solve the distribution question, we must first address the measurement problem. It is important to note that my usage of the term 'measurement' does not imply a commitment to a gradualist thesis of consciousness or to the possibility of mathematically quantifying the level of consciousness in a specific mental state of a certain system, as proposed by the Integrated Information Theory (Tononi & Koch 2015). Here, the term 'measurement' only refers to the ability to identify the presence of consciousness within a particular cognitive system. I will remain neutral regarding metaphysical theories of consciousness, such as gradualism or the all-or-nothing view (see Godfrey-Smith 2020).

There are at least three approaches to addressing the measurement problem: the theory approach, the epistemic approach, and the direct-perception approach (Andrews 2020; Barcellos 2022; Birch 2022). In what follows, I will provide a brief overview of each of these approaches.

The direct-perception approach claims that the ascription of consciousness to a given animal is not derived from inferences based on its behavior or expression. Rather, it is ascribed through direct perception of its supposedly conscious states. According to this viewpoint, consciousness is perceptible in animals in a way similar to a similar way to how we supposedly perceive the table in front of me.[4] Colin Allen (2004) highlights that Dale Jamieson (1998) and John Searle (1994) are among the main proponents of this approach. They argue that, whether through our interactions with animals or by perceiving their expressions, we are capable of directly perceiving the presence of consciousness within them (Jamieson 1998; Searle 1994). That is, when we observe an expression $E_n$, such as a dog crying or a chimpanzee's facial expression of surprise, we directly perceive that the animal possesses consciousness because these kinds of behaviors usually denote conscious experience. The direct-perception approach faces several challenges, but due to the scope of this paper, I will not address them here (see Allen 2004).

On the other hand, the theory approach offers a response to the measurement problem by advocating for the *implicit* or *explicit* application of a theory of human consciousness to animals. According to this approach, the investigation should begin by constructing or selecting one of the numerous theories of human conscious-

ness found in the literature. Only by judiciously selecting a theory can we examine whether other animals possess consciousness (Dennett 1995b; Lau 2022; Prinz 2018; de Weerd 2024). In this sense, whether an animal instantiates consciousness is always relative to whether it meets the criteria imposed by the given theory of consciousness in humans.

The epistemic approach claims that a commitment to a specific theory of consciousness is not necessarily required to address the measurement problem of animal consciousness (Andrews 2020; Birch 2022; Shevlin 2021; Tye 2017).[5] The epistemic approach argues that "we need not make major theoretical commitments in advance about the underlying mechanisms of consciousness" (Shevlin 2021, p.306). The most effective method for ascribing consciousness to animals would be through pre-theoretical factors, such as markers of consciousness, that are independent of any endorsement of a given theory of consciousness. The epistemic approach emphasizes the need for a methodology that *justifies* the best hypothesis regarding the presence or absence of consciousness in a given animal (Tye 2017, p.11).

Carruthers rejects the epistemic approach and commits himself to the theory approach. He dismisses the epistemic approach based on a critique of one of its views, namely Tye's (2017) theory-neutral approach. Tye defends two theses that he considers to be metaphysically modest, namely the *Simple View* and *Newton's Rule*. The Simple View, roughly speaking, claims that 1) "a creature is conscious at time *t* if and only if it is undergoing one or more experiences at *t*" (Tye, 2017, p.12) and 2) "a mental state is conscious if and only if it is an experience" (*ibid*.). Throughout Chapter 2 of his 2017 book, Tye seems to advocate the thesis that all perceptual states are conscious states. For example, he argues that the inattentive driver thought experiment does not imply the existence of *unconscious* perception. Tye asserts that the perceptual mental state $M_1$ instantiated by the driver is phenomenally conscious, though it is not accompanied by a higher-order mental state $M_2$ that makes the driver aware of $M_1$. Newton's Rule states that a human *S*, when presented with a stimulus *P*, exhibits behavior, or a cluster of behaviors, *B* caused by a conscious mental state *Q*. Then, if there is *no defeater*, we should consider that the best explanation for an animal *R* exhibiting behavior, or a cluster of behaviors, *B'* in response to a stimulus *P'* is the causation of *B'* by a conscious mental state *Q'*. Therefore, *R* would instantiate a conscious state *Q'* similar to the conscious state *Q* of the human being. For instance, when humans encounter a noxious stimulus, such as a rose thorn prick, we typically experience a conscious mental state like *pain*, which causes behaviors such as withdrawing from the rose and protecting the injured part of the body. Similarly, if a dog finds itself in a comparable situation and exhibits similar behaviors, and if there is *no defeater*, we can ascribe the cause of the dog's behaviors to the conscious state of *pain*.

I will not detail Carruthers' argument here, but broadly speaking, he contends

that because Tye considers all perceptual states to be conscious, he cannot distinguish flexible behaviors caused by unconscious perceptual states from those caused by conscious ones. This leads Newton's Rule to incorrectly attribute consciousness to animals whose behaviors are caused by unconscious perceptions (Carruthers, 2019, Chapter 3). To resolve this issue, Carruthers argues that Tye must be able to distinguish conscious perceptual states from unconscious ones. However, this distinction can only be made by employing a theory of consciousness. Thus, Tye's epistemic position faces a dilemma that Barcellos (2022, Chapter 3) calls "*The Carruthers-Dawkins Challenge*".[6] The challenge can be stated as follows: for Newton's Rule to be successfully applied, Tye must be able to distinguish conscious from unconscious perceptual states. However, to do so, he must rely on a theory of consciousness, which would compromise the theoretical neutrality required by the epistemic approach. On the other hand, if Tye neglects this distinction, the application of Newton's Rule is doomed to fail, as it would erroneously attribute conscious perceptual states as causes of behaviors that were, in fact, caused by unconscious perceptual states. I believe that other epistemic approaches, such as the theory-light approach (Birch 2022), can avoid this challenge. However, due to the scope of this paper, I will not advocate for any specific epistemic approach.[7]

The following sections will delve into the nihilist view involved in Carruthers' theory approach, particularly as presented in his more recent works (2018a, 2018b, 2019, 2020).[8] Carruthers (2019, Chapter 7) argues that there is *no fact of the matter* regarding the presence of consciousness in other animals. Consequently, Carruthers suggests that both the fields of science and philosophy of consciousness should abandon research on animals, since it is a stillborn project. In what follows, I will examine the motivations and arguments put forth by Carruthers and attempt to argue that his positions contain internal contradictions.[9]

## 2. Consciousness as an all-or-nothing phenomenon

With regard to its nature, Carruthers (2019, Chapter 1) contends that consciousness is an all-or-nothing phenomenon. Roughly speaking, this account supports the following disjunction: given a mental state $M_1$ instantiated by a cognitive system $S$, either $M_1$ is conscious, or $M_1$ is not conscious. In other words, $M_1$ is conscious if and only if *there is definitely something it is like to be in it*. Consciousness would be a *binary discrete phenomenon*, where it is definitely present, or definitely absent. Consciousness, therefore, would not admit degrees or borderline cases in which the presence or absence of conscious states is indeterminate, as certain gradualist theories seem to advocate (Birch et al. 2020; Godfrey-Smith 2020; Lee 2022).[10]

For instance, consider the property of being bald. Some individuals are clearly

bald, while others clearly have hair. However, between these two extremes, there exists a range of possible degrees where individuals can be situated. Moreover, within the scale of baldness, there is a point where some individuals may fall into a kind of 'limbo', where they could be considered neither fully bald nor distinctly hairy. This seems to demonstrate that baldness is a vague property, as it allows for borderline cases where it is indeterminate whether an entity *P* instantiates or does not instantiate the property *Q*. Let's shift our focus to the concept 'square'. Unlike 'baldness', 'square' seems not to be a vague concept, because it seems not to be possible to conceive a figure that resembles a square but is not a square. A geometric figure is either square or not. Therefore, there is no limbo regarding being a square (Michel 2021, p.620).

In a similar way, Carruthers (2019) regards phenomenal consciousness as an entity analogous to that of a square, i.e., just as a geometric figure is either definitely a square or not, an individual either bears consciousness or does not. He also argues against the existence of borderline cases of consciousness and degrees of consciousness. According to him, because it is challenging to conceive of scenarios where an individual is partially conscious, consciousness must be an all-or-nothing phenomenon (Carruthers, 2018b, p.53).[11]

However, Carruthers (2019) acknowledges that there can be degrees concerning the *content* of a conscious state. It is intuitively understandable that, according to him, one can be more or less conscious of a specific stimulus. For instance, I might be more or less conscious of the redness of the book in front of me, or more or less conscious of the sound coming from my window. But the variation of degrees pertains to what I am conscious *of* — i.e. what David Rosenthal (1993) calls *transitive creature consciousness* — rather than the conscious state itself. Regardless of the richness of the perceived stimulus, one is still either experiencing it or not. There is always a definite aspect of consciously noticing the redness of the book or hearing the sound near the window. In other words, irrespective of the varying content within that conscious state, the experience itself is either present or absent.

But consciousness does seem to exhibit degrees when an individual is in the process of waking up or falling asleep.During the waking-up phase, consciousness tends to increase in richness, while during the falling-asleep phase, it seems to fade. However, once again, it is not phenomenal consciousness itself that allows for degrees, but rather '*intransitive creature consciousness*' (Rosenthal 1993). This refers to the states of being awake compared to being asleep or in a coma. In both the aforementioned cases, the individual would possess consciousness in the phenomenal sense, as there is a definite subjective experience akin to waking up or falling asleep. As Carruthers claims, "no matter how impoverished their contents, it will be determinately *like something* to be in them." (2019, p.22). Thus, Carruthers concludes that phenomenal consciousness is an all-or-nothing phenomenon.

But what properties are involved in consciousness? Due to numerous anti- physi-

calist arguments — such as the explanatory gap, Mary's room, and the conceivability of zombies — some authors have argued that consciousness would involve special properties, the so-called '*qualia*' (Chalmers 1996; Jackson 1982; Levine 1983). Carruthers (2018a, 2019) calls '*qualia realism*' the view that defends the existence of such special properties. According to it, qualia would be properties that are intrinsic and irreducible to any physical state, whether representational, functional, or neuronal. Furthermore, they would be private to the subject of the experiences. That is, the peculiar experience of a certain conscious state *Q* would be restricted only to the subject who is experiencing *Q*. Despite qualia realism being compatible with certain physicalist theories, Carruthers argues that it usually involves a tacit commitment to some kind of dualism. The reason is because qualia realists "think that all physical facts can be fixed, and yet facts about qualia can vary or be absent altogether." (Carruthers 2019, p.11).

According to qualia realism, it is the intrinsic and irreducible characteristics of qualia that grant pain its unpleasant feeling. This occurs when qualia become connected to the physical state that corresponds to pain. For instance, when my injured hand is exposed to a noxious stimulus, it triggers the activation of specific neurons in the parietal lobe. According to qualia realism, the qualitative aspect of pain that I experience is not solely granted by this specific neuronal activity, but rather through the association of qualia with this neural ensemble. In other words, in order for a subject *S* to experience a qualitative state *Q*, it is insufficient for *S* to be undergoing a specific neuronal activity *N*, as it is the qualia properties *R* linked to *N* that confer *Q* to *S*.

Carruthers (2019, p.13-16) rejects qualia realism and embraces what he terms '*qualia irrealism*'. Qualia irrealism encompasses various interpretations, all sharing the common goal of associating consciousness with some kind of physical or natural property. Unlike qualia realism, qualia irrealism embraces reductionism in order to reduce consciousness to physical or physically realized properties. According to this view, there are no special, extra, or ineffable properties intrinsic to experience itself, and these properties are not irreducible to physical properties and processes. Thus, according to Carruthers (2020), conscious states can be reducible to, e.g., representations that enter the global workspace and are widely shared among consumer systems, such as rational action, verbal reporting, and metacognition.

Although Carruthers (2019, p.14-15) notes a close relation between qualia irrealism and illusionism, he rejects applying the latter label to his theory. Illusionism, roughly speaking, claims that phenomenal properties do not exist, and that the qualitative aspect of conscious experience is an *illusion* generated by internal physical mechanisms, such as introspection (Frankish 2016) and our simplified self-model that infers the cause of incoming signals from the world and our body (Clark 2023). Carruthers rejects this characterization because illusions, according to him,

arise spontaneously, whereas the belief in the existence of qualia depends on the explicit articulation of thought and the construction of philosophical arguments. That is, while standard illusionism considers internal mechanisms as the cause of our belief that phenomenal properties exist, Carruthers considers external causes, such as the articulation of arguments and modes of reflection specific to a certain scientific-philosophical community, as the cause of such beliefs.[12]

Qualia irrealism is not immune to the challenges presented by anti-physicalist arguments. For instance, it is *prima facie* conceivable that a philosophical zombie could have representations entering the global workspace without being accompanied by any qualitative experience. To address these challenges, Carruthers employs the 'phenomenal concept strategy'. He argues that combining the global workspace theory with the phenomenal concept strategy provides a correct and fully reductionist explanation of consciousness, at least in the human case. In the following two sections, we will briefly explore Carruthers' approach to these theories.

## 3. The Global Workspace Theory

For many authors, the global workspace theory (GWT) is regarded as one of the major theories of human consciousness (Block 2009; Mylopoulos 2022; Seth & Bayne 2022).[13] GWT claims that a mental state achieves consciousness when it enters the global workspace and is widely shared among various consumer systems. In other words, when information enters the global workspace and is shared across systems such as reasoning, planning, decision-making, working memory, and guided action, it becomes a conscious state (see Figure 1). Consciousness, according to GWT, would be "*just* brain-wide information sharing" (Dehaene 2014, p.165, emphasis added). On the other hand, unconscious mental states are those that fail to enter the global workspace and, consequently, are not widely shared within the brain.

Contrary to certain interpretations of GWT, Carruthers does not directly reduce or identify consciousness with specific neural patterns. Although Carruthers incorporates neuroscientific evidence to support his theory, he primarily describes GWT in cognitive terms. According to Carruthers, two reasons explain why conscious states should not be explained solely by brain processes. First, there are various explanatory levels between brain facts and consciousness facts, such as functional, computational, and representational explanations. Second, individuals commonly hold the intuition that mental states and physical states are ontologically distinct, which is compatible with the traditional mind-body problem.[14] The argument can be summarized as follows: if there are multiple levels of potential explanations bridging brain facts and consciousness facts, and if people strongly hold intuitions regarding an ontological distinction between mental and physical states, then it is not immediately necessary to explain consciousness purely in brain's vocabulary.

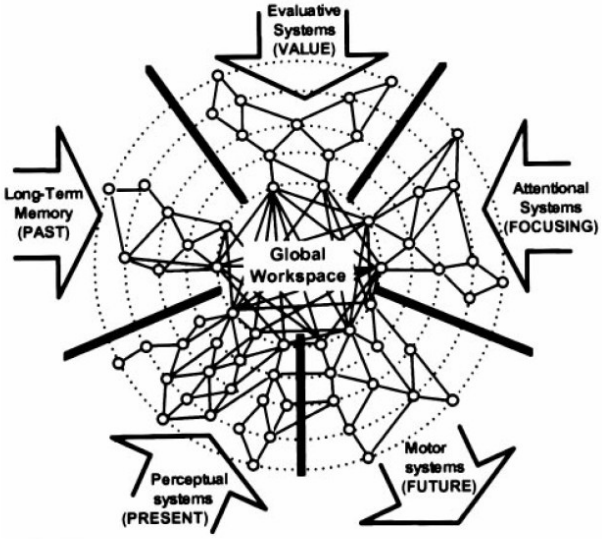


Figure 1: The Global Workspace schema illustrated. Dehaene et al. 1998, p.14530. © 1998 National Academy of Sciences.

Carruthers states that the most effective approach to explain consciousness and avoid people's dualistic intuitions is by employing *mental terms*. From this perspective, consciousness could be reduced to intentional contents that fulfill specific causal functions. The reduction of consciousness would involve providing a comprehensive explanation solely in intentional terms. However, some might argue that following a naturalistic approach would entail reducing intentional content to lower-level physical states, such as specific neuronal patterns, or even molecular and cellular processes (for the latter, see Bickle 2003 and Reber et al. 2023). This is not Carruthers' proposal. Instead, Carruthers argues that representations and representational contents are genuine psychological properties that have already been fully integrated into the naturalistic framework of the mind sciences, particularly cognitive sciences and psychology. Similar to Tyler Burge's (2010, 2022) view, Carruthers (2019, p.116-120) adopts a primitivist standpoint regarding representation and content.

## 4. The Phenomenal Concept Strategy

In the previous section, we saw that Carruthers puts forth the view that GWT offers a comprehensive explanation of consciousness from both a reductionist and naturalistic perspective. The author contends that GWT should not be regarded as providing a mere correlational description between consciousness and global broadcasting of information. Instead, Carruthers suggests that GWT should encompass a third-person

psychological framework capable of explaining all aspects of consciousness using only mental terms. For this reductionist approach to be successful, Carruthers emphasizes the need to address arguments from anti-physicalists. He argues that the dissolution of these arguments lies in the phenomenal concept strategy (PCS), which will supposedly pave the way for the triumph of his naturalistic endeavor.

The PCS states that anti-physicalist arguments, such as Mary's room and the explanatory gap, arise:

> [N]ot because our conscious states possess special properties of any sort (qualia), but rather from the distinctive first-person way in which it is possible for us to *think about* those states. It is because these thoughts lack conceptual connections with physical, functional, or even representational ones that we can conceive of the possibility of zombies; and it is also why we can see that even a complete physical/functional/intentional description of ourselves wouldn't entail the truth of those thoughts. But the nature and existence of the thoughts themselves can be explained in a way that doesn't introduce any special nonphysical properties, either. (Carruthers 2019, p.121).

As stated, PCS rejects the existence of an ontological gap between physical and non-physical properties, because reality would be entirely composed of physical properties and entities. However, PCS acknowledges an epistemic gap because our comprehension of phenomenal properties relies on dualistic vocabulary. Carruthers (2019, p.123–125) argues that phenomenal concepts are somewhat separate from third-persons concepts — such as physical, representational, and functional concepts. Therefore, according to him, we can refer to our subjective conscious experience, such as a painful experience, without invoking third-person concepts. Despite being distinct from third-person concepts, phenomenal concepts still have physical properties as their referents. More precisely, according to Carruthers (2020, p.8-9), phenomenal concepts refer only to globally shared non-conceptual content.

Regarding the nature of phenomenal concepts, Carruthers advocates the idea that phenomenal concepts are acquaintance-based indexicals, derived from globally shared non-conceptual content (see also Prinz 2012). To illustrate, consider that I am currently perceiving the blue color of the glass in front of me under normal lighting conditions. The information about this blue color, or a significant portion of it, reaches my retina, is processed by my visual system, and is then globally shared throughout the brain. This process generates the perceptual experience of blue (I will exclude the glass content for simplicity).[15] Once this non-conceptual content, namely the blue color, enters the global workspace and becomes widely shared, I can utilize it for various purposes, such as reporting that the blue glass is beautiful. Furthermore, as this perceptual content enters my global workspace, I gain the ability to form a new concept in a unique manner: the phenomenal concept 'blue experience'. According to Carruthers, the phenomenal concept that arises from my direct experience of

blue takes the form of an *indexical*, such as '*this-R*' where '*R*' means a phenomenal experience of some sort. I will not address whether this approach is successful as a response to anti-physicalist arguments — see Birch (2020) and Carruthers (2019, Chapter 6).

But how does the combination of GWT with PCS lead to Carruthers' nihilistic stance on animal consciousness? In the next section, I will summarize his argument.

## 5. Motivating the nihilistic stance towards animal consciousness

Jonathan Birch (2020) succinctly summarizes what motivates Carruthers' nihilistic stance on animal consciousness in what he calls the *'surprising conditional'*:[16]

> If GWT is a correct and fully reductive theory of phenomenal consciousness in humans, then, for all non-humans, there is no fact of the matter as to whether or not they possess phenomenally conscious states. (Birch 2020, p.7)

As Birch (2020, p.7-8) points out, Carruthers' conditional consists of the following *modus ponens*:

*Premise 1:* GWT is a correct and fully reductive phenomenal consciousness in humans.

*Premise 2:* If GWT is a correct and fully reductive theory of phenomenal consciousness in humans, then there is no fact of the matter as to whether non-human animals instantiate phenomenally conscious states.

---

*Conclusion:* For all non-human animals, there is no fact of the matter as to whether or not they instantiate phenomenally conscious states.

What motivates the radical argument presented above? Carruthers (2019, p.143) claims that it arises from the *mismatch* between the all-or-nothing nature of consciousness and the all-or-nothing nature of the human global workspace, versus the gradualist nature of the animal global workspace. According to what has been presented, if GWT is supposedly the correct and completely reductionist theory of consciousness in humans, then consciousness is all-or-nothing.

However, when exploring the cognitive and neuronal architectures of other animals, distinct forms of global workspace become evident (albeit they may bear some resemblance to those found in humans). Notably, no known animal on our planet possesses a consumer system such as verbal reporting, and only a few exhibit metacognition, planning, reasoning, and theory of mind (Andrews 2020; de Waal 2016; Wynne

& Udell 2021). Even the consumer systems shared between humans and animals display variations in complexity. For instance, our capacity for introspection appears to be more complex than that of our closest phylogenetic relatives, such as chimpanzees, and our distant relatives, such as cuttlefish. This is not to say that chimpanzees and cuttlefish lack complex cognition. The evidence available points to the opposite conclusion — that both species seem to have a complex cognitive economy (Schnell et al. 2021; Tomasello 2023). In summary, what we observe across different lineages are diverse configurations of the global workspace — some more akin to those of humans (e.g., great apes), while others are entirely distinct (e.g., cephalopods).

As it appears, there is a mismatch between consciousness, which is an all-or-nothing phenomenon and is completely explained by GWT in humans, and the global workspace of other animals, which admits degrees. Consequently, once we reject the gradualist thesis, we cannot assert that animals have different levels of consciousness based on their cognitive and neuronal architecture. Thus, we are left with two alternatives: either we acknowledge that the states *E* exhibited by some animals are conscious, or that *E* does not constitute a conscious state, as any other stance would lead to a contradiction (Carruthers 2019, p.143). This mismatch can be outlined as follows:

> *Argument 1:*
>
> Consciousness is either all-or-nothing or gradual.
>
> If consciousness is all-or-nothing, then it is not a gradual phenomenon.
>
> Consciousness is all-or-nothing.
>
> ---
>
> So, consciousness is not a gradual phenomenon.

Assuming that consciousness is just globally broadcasting nonconceptual content:

> *Argument 2:*
>
> Consciousness is an all-or-nothing phenomenon.
>
> In humans the global workspace is an all-or-nothing phenomenon.
>
> In non-human animals the global workspace admits degrees.
>
> ---
>
> Therefore, consciousness is an all-or-nothing phenomenon in humans and a gradual phenomenon in non-human animals.

As should be clear, there is a contradiction between Argument 1 and Argument 2.

Carruthers (2019, p.143-144) presents two potential responses to this discrepancy. The first suggests that we should reject GWT. According to this view, the existence of the mismatch implies that GWT, as a fully reductionist theory of consciousness, may not hold as true. If indeed GWT were the complete reductionist theory of consciousness, then there should not be a mismatch between the explanation of consciousness among humans and animals. As a result, it is argued that we must reject GWT. However, Carruthers insists that this is not the best option because "there isn't any mismatch between our conception of phenomenal consciousness and global broadcasting in the human case." (Carruthers 2019, p.143). According to Carruthers, GWT successfully explains all aspects of consciousness in humans, while remaining consistent with our first-person use of phenomenal concepts and 'consciousness' (Carruthers 2019, p.143-144). Moreover, the combination of GWT with PCS effectively addresses all objections from anti-physicalist arguments, serving as compelling evidence in support of GWT as the complete theory of consciousness.

The second response, which Carruthers supports, claims that there is no fact of the matter regarding animal consciousness. According to Carruthers, apart from humans, all judgments that attribute consciousness to other beings would fall under the category of '*semantic indeterminacy*'. In the subsequent section, I will examine the two arguments that Carruthers presents in support of this idea.

## 6. The negative and positive semantic arguments

To begin with, let us consider the negative semantic argument. This argument can be summarized as follows: When we employ a phenomenal concept, there is nothing within our *intentions to use it* that would indicate the degree of similarity or dissimilarity between the cognitive architecture of humans and animals that would justify claims about animal consciousness. In other words, our first-person use of phenomenal concepts does not establish their application beyond organisms whose cognitive architecture bears significant relevance to ours (Carruthers 2019, p.153). It is worth noting that Carruthers (2019, p.125) seems to endorse *intention-based semantics* (Borg 2008; Schiffer 1982). According to intention-based semantics, roughly speaking, the meaning of words and sentences $M_n$ when uttered by a subject $S$ is fixed by the psychological state $P$ of $S$, where $P$ is an intention of $S$. When applied to phenomenal concepts, intention-based semantics claims that, *mutatis mutandis*, the extension of a phenomenal concept $C_n$ is fixed by the intentions of the subject who instantiates $C_n$. By 'intentional', I refer to the deliberate and explicit articulation of the subject to refer to a conscious experience $E$ when employing $C_n$.[17]

The rationality behind Carruthers' (2019, p.125-126) endorsement of such semantics can be stated as follows. Firstly, since $C_n$ is fundamentally characterized as

being first-person, determining its extension cannot rely on the testimonies of others, such as an expert, nor can it be determined by consulting a textbook (Burge 1979). Moreover, $C_n$ is not a natural kind concept, as its reference is not grounded in any presumed underlying essence (as is the case with 'water', whose extension is fixed by the chemical compound $H_2O$).

When employing $C_n$, our sole intention is to refer to the content of our own conscious experience. Assuming *qualia irrealism*, we could say that "the extensions of one's phenomenal concepts get fixed by one's *classificatory dispositions* when employing them." (Carruthers 2019, p.125. Emphasis added). So, when one articulates a thought with the content '*this* is a pleasant experience', the extension of 'this' will be "*the set of non-conceptual contents* that would evoke the same dispositions-to-judge as those that underlie the thought in question (or something of the sort)." (Carruthers 2019, p.125-126, emphasis added).

Returning to the negative semantic argument, nothing in our use of phenomenal concepts fixes the extension to animals' mental states. For example, when I consciously perceive a blue painting under normal lighting conditions, I may have thoughts like, '*This* is a beautiful experience'. However, within this context, the acquaintance-based indexical 'this' does not by itself determine its application to cognitive architecture beyond my own. The indexical only refers to my representation of a blue paint. That is, when employing $C_n$, my intention solely fixes its extension to the first-person experience in which it arises — namely, the globally shared content among the consumer systems associated with the corresponding conscious state.

For these reasons, Carruthers argues that there exists only a spectrum of similarity between the global workspace of animals and humans. Given that consciousness does not admit degrees, a demarcation line must be drawn to distinguish conscious animals from unconscious ones as their minds become less akin to ours. However, since our use of phenomenal concepts fails to establish these boundaries, there is nothing to be discovered regarding animal consciousness. Therefore, there would be no fact of the matter involving animal consciousness. In fact, the negative semantic argument *does not* claim that *animals lack consciousness*, but rather *that there is no fact of the matter due to semantic indeterminacy*.

In contrast to the negative argument, the positive semantic argument does not rely on the assumption that consciousness is an all-or-nothing phenomenon. While it does rely on the premise that consciousness is simply globally broadcasted content, this argument focuses on the semantics of judgments ascribing consciousness to animals. Specifically, it addresses the truth conditions associated with such judgments. As Carruthers highlights, it is crucial to remember that 'consciousness' is fundamentally a first-person concept, "grounded in one's capacities for thought about one's own globally broadcast perceptual states." (Carruthers 2019, p.155).

By inquiring about the truth conditions for judgments such as "The animal *S* has

perceptual states like this-*R*" (Carruthers 2019, p.156), Carruthers outright dismisses certain options. The first of these options involves the hypothesis that our phenomenal concepts somehow refer to natural kinds. Carruthers argues against this view by claiming that when we employ a phenomenal concept, we do not do so with the intention that it represents a natural kind concept or that it tracks an underlying essence within our experience. This is a consequence of Carruthers' reliance on intention-based semantics. The second option rejected by Carruthers is that our phenomenal concepts refer to phenomenal properties. As we have seen, Carruthers denies the existence of such properties. Finally, Carruthers claims that the truth value of judgments ascribing consciousness is determined by the following counterfactual:

> Creature C has perceptual states that are *this-E* [. . . ] [i]f the dispositions-to-judge that underlie my use of the concept *THIS*-E were to be instantiated in creature C, then they would issue in a judgment that some of the creature's perceptual states are *this-E*. (Carruthers 2019, p.157)

When it comes to judgments ascribing consciousness to humans, the satisfaction of the counterfactual becomes evident. It is reasonable to assume that, under equal or similar conditions, other humans would exhibit the same dispositional judgments as I do when employing a phenomenal concept. For example, when I perceive the blue in Rothko's painting and use an indexical that refers to my pleasurable experience of perceiving its blueness, the employment of the indexical presupposes the utilization of capacities such as reflection on my thoughts concerning the painting and the language through which I express them. Hence, it seems reasonable to consider that another human is also capable of undergoing a similar color experience to what my phenomenal concept refers to (excluding individuals with visual impairments, such as color blindness). This would be supported by the shared similarities in our neurophysiology, behavior, brain ontogeny, and psychological states. Moreover, according to Carruthers, the evidence for consciousness ascription does not only rely on third-person evidence but also on our ability to conceive zombies and our ability to formulate the hard problem.

However, when it comes to animals, the counterfactual is no longer available. The reason is that assuming that an animal possesses the same dispositions to judge as mine when using a phenomenal concept would imply a completely different mind from what animals actually possess. Consider the following example taken from Carruthers (2019, p.158). Suppose I believe that monkeys have perceptual states akin to those which underlie $C_n$. In this case, I am making a judgment about the mind of the animal in question. According to the counterfactual conditional, the truth conditions underlying such a judgment would resemble the following: if the dispositions underlying my use of $C_n$ were instantiated in the monkey's mind, then these dispositions would be triggered by the monkey's own mental states. As Carruthers highlights, "the

antecedent of this conditional entails that I am no longer thinking about a monkey's mind, but the mind of another sort of creature altogether, with cognitive capacities like those of humans." (*ibid.*).

When projecting my phenomenal concept onto the minds of other animals, I would be assuming that they have a mind similar to mine. However, it is evident that the minds of most animals on Earth do not resemble mine. Thus, the counterfactual conditional that underlies ascriptive judgments is no longer available. The question of whether monkeys have conscious states similar to '*this-R*' would entail a false assumption, namely, that monkeys have a mind that is distinct from their actual cognitive architecture. As a result, the judgments that ascribe consciousness to other animals become *semantically indeterminate*. These judgments would be neither true nor false. For these reasons, Carruthers concludes that there is no fact of the matter regarding consciousness for cognitive architecture beyond our own.

In the following sections, I will present two arguments against Carruthers' view.

## 7. The unnaturalized intention-based semantics

The first argument I will present concern with Carruthers' commitment to an interpretation of *intention-based semantics* applied to phenomenal concepts. Intention-based semantics (IBS), as we have seen in the previous section, is a theory of meaning championed by the philosopher Paul Grice (1957) and further developed by other authors (see Schiffer 1972, 1982). Emma Borg (2008) distinguishes two kinds of IBS: A-style and B-style. The former maintains that the existence of an intentional agent articulating 'physical objects' (such as sounds and written words) in a community is the minimum prerequisite for a given utterance to have meaning. In other words, the meaning of a given utterance $p$ is determined by the intention of the uttering agent $S$, as agreed upon by the community $Q$ where $S$ is appropriately situated. On the other hand, B-style IBS claims that not only is the intention of $S$ a prerequisite for meaning, but it is also necessary for determining the reference of the utterance. For instance, when John says, 'Quine is the father of contemporary naturalistic epistemology', what fixes the reference of the sentence is John's *intention* — a psychological state that can be characterized by the following content: 'the American philosopher who published the paper *Epistemology Naturalized* in 1969'.[18]

Carruthers (2019, p.147-148) seems to embrace the B-style IBS, as he contends that the reference of a specific phenomenal concept is determined by the subject's *intention*. To illustrate, when I feel something as a pleasurable experience and think '*this* is a pleasurable experience', the extension of the indexical 'this' is fixed by my deliberate intentions at the moment of its usage. According to Carruthers, when we use a phenomenal concept, we only mean to refer to "the qualities we are aware of in

ourselves." (Carruthers 2019, p. 148). These qualities that we experience would be, in Carruthers' view, the globally broadcasted non-conceptual contents. In short, the reason why phenomenal concepts cannot refer to other phenomena, such as natural kinds, is that when we employ a phenomenal concept, we *intentionally refer* solely to globally broadcasted non-conceptual contents.

However, Birch (2020) highlights a dilemma arising from Carruthers' framework. As Birch argues, Carruthers' commitment to qualia irrealism leads to a contradiction when attempting to fix the reference of phenomenal concepts through the bearer's intentions. According to Birch, when we employ a phenomenal concept, we intuitively do not have globally shared information in our minds, but rather our qualitative subjective experience. This holds true regardless of the true nature of consciousness. We could even say that what we have in mind are non-physical properties, such as qualia. This may be supported by the nativist theory of dualist principles, which claims that we have an innate disposition to think of mental states as supernatural entities (Barlev & Schtulman 2021; Berent 2020, 2024; Bloom 2004). It is important to note that Birch is not suggesting that our qualitative experience is derived from phenomenal properties, whether those properties are supernatural or physical. Instead, he argues that regardless of what constitutes consciousness, when we employ a phenomenal concept, we intuitively intend to refer to the qualitative experience we instantiate. In fact, Birch provides his intuition that the intention in using a phenomenal concept is to refer to phenomenal properties. Although some authors consider intuition to be strong evidence for consciousness (Chalmers 2018; Searle 1997), others may argue that Birch's intuition could be erroneous or weak evidence. However, empirical evidence suggests that when using the term 'consciousness', people seek to refer to states with valence, such as pain and pleasure (Sytsma & Machery 2010). In this case, it seems more plausible to consider that when people use a phenomenal concept, they seek to refer to phenomenal properties such as those presented in pain, rather than globally shared information (even assuming that GWT is the correct theory of consciousness). Therefore, while the phenomenal concept may not properly track globally shared content, the crucial point is that *intention* plays the role of fixing the reference. As Carruthers himself assumes, our capacity to generate thought experiments and articulate the hard problem derives from our dualistic vocabulary, which seems to refer to phenomenal properties.

That said, Birch presents Carruthers with the following dilemma. On the one hand, if we embrace IBS, then qualia irrealism is false, since intuitively, our intention when employing a phenomenal concept $C_n$ is to refer to phenomenal properties, and what fixes the reference of $C_n$ is our intention when using it. But if we concede that our intentions do not fix the reference of $C_n$, then this opens the door to the potential for phenomenal concepts to refer to qualia, natural kinds, and other properties. Therefore, Birch argues that qualia irrealism and IBS are fundamentally incompati-

ble. To be clear, the first horn of the dilemma does not assert that the phenomenal properties people intuitively have in mind when employing $C_n$ are qualia. Making such an assertion would prematurely commit one to a philosophical thesis about the metaphysical constitution of phenomenal properties. The idea is merely that, according to available evidence from experimental philosophy and social psychology (Berent 2020, 2024; Sytsma & Machery 2010), the conception that laypeople have of phenomenal properties is not that they are widely shared non-conceptual contents in the brain. What exactly they are for laypeople, however, remains a matter of debate (with the body of evidence pointing to a dualistic view of phenomenal properties). Here we are dealing not with the metaphysics of phenomenal properties, but with the judgments that laypeople make about their constitution. Thus, it is at least evident that people are not referring to such non-conceptual contents when employing $C_n$.

The conclusion of the dilemma would be that GWT does not align with Carruthers' fully reductionist and naturalist endeavor. Carruthers could avoid this by considering that $C_n$ is non-referring, as there would be no experiential qualities. This would commit Carruthers to a radical eliminativism, something he does not seem to endorse.[19] Carruthers could also argue that $C_n$ tracks globally shared content regardless of the intention of the bearer of $C_n$. However, this would force Carruthers to reject IBS, a theory to which he himself subscribes.

But one could argue that Carruthers does not defend *intention* as the reference-fixing state of the phenomenal concept $C_n$, but rather the *disposition-to-judge* that arises when employing $C_n$ (something I introduced in the previous section).[20] This could undermine my and Birch's argument, as we consider the possibility of intentions arbitrarily fixing distinct unobservable references. While the dispositions-to-judge remain independent of my intentions when referring to $C_n$, they are therefore not arbitrary reference fixers. For instance, in an astronomy class, John might use the singular term 'the morning star' to refer to the planet Venus, while in a similar class, Terry might use 'the morning star' to refer to Phosphorus, unaware that 'Phosphorus' is Venus. Despite both terms referring to the same celestial body in our reality, the intention of each subject is different, leading to a different fixing of reference. Of course, the difference between the employment of intentions as reference-fixers in this example and in the case of phenomenal concepts is that in the former there is a method of correction that can be independently checked, whereas in the latter there is not — given that phenomenal concepts concern properties that can only be accessed by experiencing subjects.

I would like to propose two arguments to counter this objection. Firstly, from an exegetical perspective, Carruthers' use of 'disposition-to-judge' as the reference-fixer, rather than 'intentions', is ambiguous. In some texts, Carruthers explicitly states that intentions fix the reference (Carruthers 2018b, p.53; 2020). Even in his 2019 book, Carruthers, in certain passages, sometimes explains that intentions fix the reference

(2019, p.125), and other times, he claims that it is the disposition-to-judge that fixes the reference (2019, p. 157).[21] Therefore, from an exegetical point of view, it is unclear whether Carruthers upholds the disposition-to-judge, rather than intentions, as the reference fixer for a given phenomenal concept.

However, I believe that we still encounter an issue even if we assume the disposition-to-judge interpretation of Carruthers' work. Our intention can arbitrarily fix a reference, since, *prima facie*, it may be subject to the idiosyncratic articulations of the subject. For instance, when expressing the proper name 'Quine', my intention may be directed towards 'the American philosopher father of naturalized epistemology', while for John, the intention may be directed towards the creator of the character Mafalda. In this case, my intention appropriately fixes the reference of the proper name, since the proper name correctly tracks the definite description, whereas for John the relation is not appropriate.[22] But one could argue that our dispositions-to-judge are independent of the subject's idiosyncratic articulations. Nevertheless, this claim is not entirely clear to me. Even if we were to accept it as true, the problem of reference fixation persists. The network of dispositions of a given subject, when employing a phenomenal concept in a certain context, may differ from that of another subject using a similar phenomenal concept in a similar context. For instance, it is conceivable that when one subject is thinking '*this* is a good experience' in context *p*, another subject in a similar context *p'* could exhibit different dispositions than those they present. This variation may stem from differing cultural or even innate backgrounds. Carruthers (2019, p.158) contends that people's behavior, such as their capacity to grapple with the hard problem, provides compelling evidence supporting the validity of others' thoughts on consciousness. However, from an empirical standpoint, this assertion seems contentious. Some evidence suggests that people's intuitions about such cases are not so widespread and, when they do arise, they may not pertain to the nature of consciousness (Díaz 2021). Furthermore, there would be no independent method of verifying the underlying mental states that carry the reference of each subject's phenomenal concept, as these states are accessible only to the subject of experience. Indeed, Carruthers could argue that even if one concedes the validity of these semantic concerns regarding the fixation of the content of the phenomenal concept, there is still abundant empirical evidence in humans supporting GWT as the correct theory of consciousness. In summary, these semantic problems do not imply the refutation of GWT. Two points must be made clear. First, although GWT is one of the leading theories of human consciousness available, it still faces conceptual and empirical criticism (see, e.g., Block 2009; Lau 2022; Melloni et al. 2023). For these reasons, GWT is not the only game in town. Second, the argument presented here *does not aim to refute* GWT, but to demonstrate the limitations of Carruthers' arguments that there is no fact of the matter regarding animal consciousness. As Carruthers makes clear, his skeptical arguments rest on semantic issues. So,

if Carruthers were to concede the validity of the semantic concerns presented here, his skeptical argument would lose strength. In a nutshell, we don't need to refute GWT to undermine Carruthers' skeptical arguments, as the nature of his skepticism is semantic rather than metaphysical, epistemological, or scientific.

In his recent works, particularly in the book *Human and Animals Minds* published in 2019, Carruthers tries to establish a robust naturalistic philosophical framework that takes empirical evidence seriously. He even acknowledges that his shift toward first-order representationalism was driven by the lack of empirical support for higher-order representationalism (Carruthers 2019, p.x-xi). However, when it comes to the semantics that he adopts, it appears that Carruthers may not be aware that he relies primarily on his private intuitions regarding what people are presumed to think when employing a phenomenal concept. It is not entirely clear to me that when we employ a phenomenal concept, our intention is to refer exclusively to representational content being globally broadcast or even to a qualitative experience, as Birch suggests.

When an individual uses a phenomenal concept, they may intentionally seek to refer to something entirely distinct from and incompatible with what others intend to refer to when using the same or similar phenomenal concept in a similar context. More broadly, in a given context *O*, an individual *S* using a phenomenal concept $C_1$ may have in mind a reference $R_1$ that is entirely different from the reference $R_2$ held by individual *Q* when using an equal or similar phenomenal concept $C_2$ in an equal or similar context *O'*. For instance, while enjoying an ice cream on the beach, one might have the thought '*This* is a pleasurable experience'. Different individuals may offer diverse explanations regarding what they had in mind when employing the indexical. A panpsychist may argue that when using the indexical, their intention is to refer only to a certain intrinsic qualitative property underlying all physical matter. Conversely, a property dualist might contend that they aim to refer to supernatural properties. A proponent of embodied cognition might express the intent to refer to the entire body and sensorimotor lawlike contingencies — and so on.

In fact, Carruthers could argue that, as he demonstrated, GWT is the true theory of consciousness. In his 2019 book, Carruthers sought to refute some competing theories to GWT, such as the Integrated Information Theory proposed by Giulio Tononi and colleagues (Carruthers 2019, p.73-77; Tononi & Koch 2015), the Fragile Short-Term theory championed by Ned Block (Block 2011, 2023; Carruthers 2019, p.80-86), and some versions of higher-order theory (Carruthers 2000, 2019, p. 86-95; Rosenthal 2005). Even if we admitted that Carruthers indeed refutes these theories, other important theories that Carruthers does not address would remain in contention — as is the case with anti-representationalist theories (e.g., sensorimotor theories) (see O'Regan 2011; O'Regan & Noë 2001). So, sensorimotor theorists could still argue that when using a phenomenal concept, they intend to refer to certain sensorimotor interactions.

What about the so-called 'laypeople'? Carruthers could argue that when he uses the term 'we', he specifically refers to the general population who are not experts in philosophy and/or the sciences of the mind. He may exclude specialists from his sample, suggesting that they have biases due to their favorite theory of consciousness. However, this line of reasoning does not convince me. If such a claim were put forward, it would appear entirely *ad hoc*, since its sole purpose would be to support the semantics underlying Carruthers' positive and negative arguments. Furthermore, there is no guarantee that laypeople would share the same intentions and dispositions, as Carruthers claims, when employing a phenomenal concept to refer to *P* rather than *Q*. In fact, compelling evidence from experimental philosophy suggests that certain intuitions often assumed to support the plausibility of reference theories are actually limited to Western populations (Machery et al. 2004; Mallon et al. 2009). Moreover, some intuitions regarding the generation of the hard problem are not as widespread among laypeople (Díaz 2021).

Given the naturalistic spirit of Carruthers' endeavor, it would be most appropriate and parsimonious to establish a semantics based on empirical research concerning what individuals truly have in mind when employing a phenomenal concept in a certain context.[23] However, several issues remain unsolved. To ensure research outcomes are minimally reasonable, it should employ the most rigorous methodologies demanded by the relevant fields such as anthropology, experimental philosophy, linguistics, and social psychology. For instance, the research would need a sufficiently large and qualitatively diverse sample, encompassing not only populations classified as WEIRD (Western, Educated, Industrialized, Rich, and Democratic societies) but also non-WEIRD populations, such as hunter-gatherer communities (Henrich 2021; Henrich et al. 2010). Even if such research were successfully conducted and revealed that most people indeed have qualities in mind equal to those proposed by Carruthers, there would likely still be individuals who claim to have different intentions when using the same or similar phenomenal concept in a similar context. We might even encounter people asserting that when they employ a phenomenal concept, they seek to refer simultaneously to both experiential qualities and natural kinds, as suggested by Bayne and Shea (2020). In fact, through the study of the English corpus, researchers have discovered that when using the term 'experience', both academic and non-academic individuals seek to refer to distinct things. Moreover, even within the philosophical community itself, the concept is employed to refer to distinct things. Notably, it is only within the domain of philosophy of mind that the term is predominantly employed to refer to mental states, such as conscious states (Sytsma & Fischer 2023). Other evidence also suggests that laypeople and experts differ on their folk notion of 'subjective experience' (Sytsma & Machery 2010; Sytsma & Ozdemir 2019). So, it appears plausible that people with different backgrounds would refer to different things when employing a phenomenal concept. It remains unclear to me

how Carruthers could reconcile his armchair intuitions with the actual intentions of individuals revealed by empirical evidence.

Furthermore, Carruthers does not provide sufficient justification for the adoption of intention-based semantics. Why should we accept this semantic approach over another that is empirically motivated and more apt to address the problem at hand? The issue of semantic indeterminacy may stem from the underlying semantics (IBS), rather than from a unique characteristic of the area of animal consciousness. In summary, it is possible that Carruthers' radical conclusion merely exposes the limitations of his chosen semantics.

## 8. Is consciousness all-or-nothing?

Throughout his recent writings, Carruthers claims that consciousness is an all-or-nothing phenomenon. However, this thesis does not seem to be self-evident. In what follows, I will present an argument highlighting the incompatibility between Carruthers' reductionist version of GWT and the thesis that consciousness is all-or-nothing. The structure of my arguments can be summarized as follows:

*Argument I:*

1. Consciousness is either all-or-nothing or gradual.
2. If consciousness is physical and all-or-nothing, then it is not a gradual phenomenon.
3. Consciousness is physical and all-or-nothing.

---

4. So, consciousness is not a gradual phenomenon.

*Argument II:*

1. Consciousness consists only of nonconceptual representational content globally broadcasted.
2. Representational content is a physical phenomenon that admits degrees.

---

3. Therefore, consciousness is a physical phenomenon that admits degrees.

From arguments I and II, we can conclude that consciousness is a physical phenomenon that is all-or-nothing and gradual at the same time, which presents a contradiction. In what follows I will examine the main points of Argument I and then Argument II.

Premise (B) of the first argument asserts that if consciousness is both a physical phenomenon and an all-or-nothing phenomenon, as Carruthers claims, it cannot simultaneously be a gradual phenomenon. This premise acknowledges two potential categories of physical phenomena into which consciousness could fall: all-or-nothing and gradual. Some authors, notably Michael Tye (2021), argue that nearly all physical phenomena are gradual, except for those within a hypothetical *micro-reality* — an elusive realm beyond even quantum reality, purported to encompass intrinsic properties such as quiddities. However, for the sake of the argument, I will set aside the existence of this enigmatic realm proposed by Tye and other panpsychists. What is crucial in premise (B) is the recognition of the existence of physical phenomena that are either all-or-nothing or gradual. It asserts that if consciousness is categorized as being all-or-nothing, then it cannot simultaneously exhibit degrees.

Conclusion (D) logically follows from the premises, as if consciousness is an all-or-nothing physical phenomenon, it cannot be its opposite (i.e., a gradual phenomenon). It is noteworthy that Argument I does not present a contentious point within Carruthers' proposal; rather, it merely encapsulates his argument.

Now let's examine the premises of *Argument II*. Premise (A) reflects Carruthers' commitment to qualia irrealism and GWT. As discussed earlier, Carruthers (2019, p.116-120) asserts that phenomenal consciousness must be fully explained in a reductionist fashion, using only representational terms. Moreover, Carruthers (2019, p.118-119) argues that representational content and representations do not require reduction to lower-level entities, such as those in physics or chemistry. Instead, he considers them to be primitive entities, inherently irreducible to lower-level properties. Carruthers provides several reasons to support this claim, summarized as follows. The first argument posits that (I) representational content is a fundamental postulate in cognitive psychology, (II) psychological properties and the integration of psychology and neuroscience are real, and (III) therefore, it is reasonable to assume that representational contents are also real. The second argument asserts that representational contents and representations are already considered '*acceptably naturalized*' properties, playing a crucial role in cognitive science, psychology, and neuroscience (Carruthers 2017, p.75). Hence, Carruthers maintains that representational contents and representations are primitive entities that have already been naturalized.

This statement is rather contentious. There is a growing movement of authors advocating for the elimination of mental notions such as representation and content in favor of explanations rooted in, e.g., dynamical systems (see Beer & Beer 2000; Chemero 2000, 2009; Favela 2024; Hutto & Myin 2013, 2017). At no point in his recent texts on animal consciousness, especially in his 2019 book, does Carruthers provide a clear justification for the appeal to representation. Instead, Carruthers claims vaguely that psychology and neuroscience largely use representation in their expla-

nations, and therefore, we must be realistic about them. However, a recent survey regarding the perspectives of psychologists, neuroscientists, and philosophers ($N = 736$) on the concept of 'representation' revealed that while participants generally agree that representation involves the conveyance of information, they showed a preference for non-representational explanations of brain activity (Favela & Machery 2023). Nevertheless, it is important to note that this does not imply that representations do not exist or lack explanatory significance. The survey could align with theories that claim that a naturalistic explanation of representation is successful only when the natural properties treated as representational actually play a role in genuine representational explanations in the sciences of the mind (see Neander, 2017; Shea 2018). For instance, these natural properties may be certain neural firings in the primary visual cortex responding to distal properties of the environment such as occluded scenes (Papale et al. 2023). Nonetheless, contrary to Carruthers' assertions, these naturalistic theories are reductionist. They seek to reduce representational properties to natural properties such as neuronal activations or biological functions. Therefore, *prima facie*, Carruthers cannot appeal to this approach to align his primitivist theory of content with the results of the survey presented.

Premise (B) asserts that representational contents are physical phenomena that admit degrees. Let's consider two examples taken from Tye (2021, p.7) to illustrate this point. Take the property of representing 'meat'. Over a certain historical period, 'meat' may have initially meant something like '*being edible*'. As time passed, the meaning of 'meat' evolved to refer specifically to 'being flesh', as we understand it today. Claiming that this change in meaning was abrupt, transitioning from 0 to 1, appears implausible. It is more likely that the change occurred gradually over time and in different geographic regions, such that at certain points it was unclear whether the concept definitely meant 'being edible' or 'being flesh'. For instance, the change may have initially taken place in London and later in Carlisle, leading to borderline cases during the transition. As Tye rightly points out, the property of representing something allows for borderline cases, indicating that the very nature of representing something admits degrees.

Two objections could be raised. The first is that one might argue that the transition from meaning $x$ to $y$ of $p$ could be gradual at the community level but abrupt at the individual level. For example, the change in the meaning of '*meat*' from '*being edible*' to '*being flesh*' could have been *statistically gradual* for the *population* of London and Carlisle. But the change in meaning would have been *abrupt* at the *individual level* — e.g., for John, an adult resident of Carlisle, the change occurred immediately. Although this is metaphysically possible, I believe it is empirically implausible. Generally, when we are compelled to represent a state of affairs differently, this does not occur abruptly at the personal level. Suppose that John represents *meat* as *being edible* at time $t_1$. Also, consider that due to contextual contingencies, the "represen-

tational novelty" of *'meat'* meaning *'being flesh'* reaches Carlisle at $t_2$. What would likely occur at $t_2$ is a gradual adjustment by John to represent *meat* as meaning *'being flesh'* rather than *'being edible'*, given that throughout his life John has always represented *meat* as *'being edible'*. In a conversation with his friend Thomas, a traveling merchant and long-time resident of London, John could even represent *meat* as meaning *'being edible'* and be corrected by Thomas. Think about how difficult it is to change your representations of something even when you consciously strive to do so. Nevertheless, someone might insist that for Ivy, an infant resident of Carlisle at $t_1$, the representational change from *'being edible'* to *'being flesh'* might occur abruptly at $t_2$. One could even argue that this would happen more easily for Ivy due to her brain possessing a greater neuroplasticity, which is capable of supporting quicker and more abrupt representational changes compared to John's mature brain. But even if we concede that this change may be abrupt for Ivy, we must consider that, according to Carruthers (2019), her global workspace differs from that of adults precisely because she is an infant and does not yet have a fully mature brain. Since GWT is a theory concerning consciousness in neurotypical adult humans, using an argument based on infants to assert that representational change at the individual level occurs abruptly would seem contradictory and *ad hoc*.

The second objection is that the example pertains to conceptual representation and, therefore, does not impact Carruthers' thesis, which identifies consciousness with certain nonconceptual contents. However, there are instances of nonconceptual contents that admit gradations. Let us briefly examine them.

To illustrate further, consider a specific set of neurons *N* firing in the cortex V1 as they represent a line *O* in the visual field of a neurotypical individual. According to a standard conception of representation, it is claimed that representing involves tracking *O* under normal conditions (Tye 2021, p.7). Given that the concept of 'normality' itself allows for degrees, the notion of representation should also admit degrees, whether 'normality' is understood from the perspective of "statistically (what t*ypically* happens)" (*ibid.*) or "teleologically (what is *supposed* to happen if the relevant system is operating properly)" (*ibid.*). Carruthers might reject such an example because he does not reduce consciousness to neuronal states but rather to cognitive states. However, consider the examples of ambiguous bistable figure stimuli (see Figure 2). During the presentation of ambiguous bistable figures, subjects generally perceive the figure in one way (A) and, after some, time perceive it in another way (B). The transition from the perceptual representation A to B occurs gradually. Interestingly, this representation shift does not necessarily involve conceptual content/abilities, since even infants and animals lacking conceptual capabilities are subject to such ambiguity (Block 2023).

Moreover, as pointed out by Souza Filho (2022), it could be argued that, similar to biological categories, representational contents and states also exhibit degrees.

Figure 2: The famous Rat-Man ambiguous figure. Bugelski & Alampay, 1961, p.206.

Consider a biological trait *T*, such as a bird's wing, and assume an etiological conception of biological function. According to the latter, a biological trait *T* has the function *F* if and only if it has been selected, through natural selection, to bring about a specific effect *E* (Neander 1991). As per the theory of evolution, the emergence of trait *T* occurs gradually through the accumulation of small and diverse genetic mutations over multiple generations of organism populations *O*, within a certain time period *Q* (Dawkins 1986; Futuyma & Kirkpatrick 2022). For this reason, the development of the biological trait 'wing' did not occur in a sudden burst but rather through the successive accumulation of mutations that gradually became more prevalent in a specific population due to their beneficial effects on fitness, thus being favored by natural selection. The ancestral form of a bird's wing probably did not possess the function of flight initially; it may have served purposes such as gliding, courtship, or thermoregulation. Over generations, with genetic mutations spreading through some populations favored by natural selection, the 'wing' trait finally emerged and acquired its capacity for flight. During evolutionary history, there were instances of traits that were neither fully developed as wings nor significantly different from them. Thus, borderline cases of wings existed, since it was unclear whether an individual possessed a wing or not.[24]

Now, if representational states and contents are also subject to the universal acid of Darwinism (Dennett 1995a), as is probably the case, then representational states and contents probably emerged through successive modifications across generations, with functions different from their current ones. Similar to biological traits, there would be a gray area where the determination of certain mental states *M* as representational and the categorization of systems *S* as representational were indeterminate. In other words, there were borderline cases of representational systems and states

(Souza Filho 2022). While this objection may seem broad, I contend that it directly challenges Carruthers' view, since he endorses a teleosemantic account of representation, which is in turn susceptible to the challenges posed by the theory of evolution by natural selection (Carruthers 2024, p. 96). The particular point of interest here is the recognition that representation is an entity that allows borderline cases.[25] It appears that biological categories are ontologically vague; therefore, representations, being biological categories, are also ontologically vague. So, given Carruthers' (*ibid*.) endorsement of teleosemantics, it follows that he should also admit that representations are ontologically vague.

However, Carruthers might concede that representations are vague entities but argue that the non-conceptual representational content of phenomenal consciousness does not admit degrees. In other words, he could agree that representational contents admit degrees but insist that those concerning phenomenal consciousness do not. But why, in the case of consciousness alone, would representations not admit degrees? The thesis that representations involved in mental states such as perception, cognition, and affectivity admit degrees, while those involved in phenomenal consciousness do not, seems completely *ad hoc*. I see no reason to uphold this thesis other than to prevent it from falling into the contradiction presented here.

Finally, conclusion (C) logically follows from the premises: If consciousness is merely globally broadcast representational content and the nature of such content admits degrees, then consciousness also admits degrees. However, as demonstrated in Argument I, consciousness cannot be a gradual phenomenon. This contradiction arises between the assertion that consciousness is all-or-nothing (Argument I) and the proposition that it is gradual (Argument II). Therefore, the overall conclusion (D) reveals the inconsistency within Carruthers' view. It is important to note that the argument does not hinge on whether consciousness is indeed all-or-nothing, but rather on the contradiction that emerges from the view that consciousness is both all-or-nothing and reducible to representational content, even though the latter admits degrees.

It is unclear how Carruthers can evade this contradiction since he establishes a reduction relationship between consciousness and representational content. Furthermore, Carruthers cannot resort to other types of relationships such as grounding or supervenience. For instance, he cannot argue that consciousness, being an all-or-nothing phenomenon, is metaphysically grounded in representational content that admits degrees. This is primarily due to two reasons that we previously discussed: (I) the establishment of a reduction relationship between consciousness and representational content, and (II) the commitment to the primitivist thesis regarding content. Hence, if Carruthers upholds the all-or-nothing thesis of consciousness, then he cannot reduce consciousness to representational content, as the latter admits degrees. This would ultimately undermine Carruthers' fully reductionist approach, rendering

his arguments that there is no fact of the matter concerning animal consciousness untenable. After all, it appears that Carruthers wants to have his cake and eat it too.

## 9. Conclusion

In this paper, I argued that Carruthers' radical position not only lacks reasonableness but also exhibits internal inconsistencies. As I have attempted to demonstrate, these inconsistencies undermine the feasibility of his reductionist and naturalist project concerning consciousness. Carruthers' project, therefore, fails to eliminate the possibility of genuine science and philosophy of animal consciousness. Given the arguments presented in this article, I believe that the science and philosophy of animal consciousness must persevere in their endeavors. Despite the controversial rhetoric, radical criticisms against this area should be approached with *caution*, with the primary goal of enhancing the techniques and conceptual clarity involved in the discussion.

## References

Allen, C. 2004. Animal Pain. *Noûs* 38(4): 617–643.

Allen, C.; Trestman, M. Animal Consciousness. In: E. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*. Winter 2016 edition. https://plato.stanford.edu/archives/win2020/entries/consciousness-animal. Access 31.12.2024.

Andrews, K. 2020. *The Animal Mind: An Introduction to the Philosophy of Animal Cognition*. 2nd ed. New York: Routledge.

Andrews, K. 2024. "All animals are conscious": Shifting The Null Hypothesis in Consciousness science. *Mind & Language* 39(3): 415–433.

Baars, B. 1988. *A Cognitive Theory of Consciousness*. Cambridge: Cambridge University Press.

Barcellos, V. M. 2022. *Como Atribuir Consciência aos Animais*. Master's Dissertation in Philosophy. Instituto de Filosofia e Ciências Sociais, Universidade Federal do Rio de Janeiro. Rio de Janeiro, p. 186.

Barlev, M.; Schtulman, A. 2021. Minds, Bodies, Spirits, and Gods: Does Widespread Belief in Disembodied Beings Imply That We Are Inherent Dualists? *Psychological Review* 128(6): 1007–1021.

Bayne, T.; Shea, N. 2020. Consciousness, Concepts, and Natural Kinds. *Philosophical Topics* 48(1): 65–84.

Beer, R. D. 2000. Dynamical Approaches to Cognitive Science. *Trends in Cognitive Science* 4(3): 91–99.

Berent, I. 2020. *The Blind Storyteller: How We Reason About Human Nature*. New York: Oxford University Press.

Berent, I. 2024. Consciousness Isn't "Hard" – It's Human Psychology That Makes It So. *Neuroscience of Consciousness* 2024(1).

Bickle, J. 2003. *Philosophy and Neuroscience: A Ruthlessly Reductive Account*. Dordrecht: Springer Dordrecht.

Birch, J. 2020. Global Workspace Theory and Animal Consciousness. *Philosophical Topics* 48(1): 21–38. [Accessed version: https://philpapers.org/archive/BIRGWT.pdf]

Birch, J. 2021. The Hatching of Consciousness. *History and Philosophy of the Life Sciences* 43(121): 1–12.

Birch, J. 2022. The Search for Invertebrate Consciousness. *Noûs* 56(1): 133–153.

Birch, J.; Schnell, A. K.; Clayton, N. S. 2020. Dimensions of Animal Consciousness. *Trends in Cognitive Sciences* 24(10): 789–801.

Block, N. 1995. On a Confusion About a Function of Consciousness. *Behavioral and Brain Sciences* 18(2): 227–247.

Block, N. 2009. Comparing the Major Theories of Consciousness. In M. S. Gazzaniga (Ed.), *The Cognitive Neurosciences*, p.1111–1122. 4th ed. Cambridge, MA: MIT Press.

Block, N. 2011. Perceptual Consciousness Overflows Cognitive Access. *Trends in Cognitive Science* 15(12):567–575.

Block, N. 2023. *The Border Between Seeing and Thinking*. New York: Oxford University Press.

Bloom, P. 2004. *Descartes' Baby: How the Science of Child Development Explains What Makes Us Humans*. New York: Basic Books.

Borg, E. 2008. Intention-Based Semantics. In: E. L. Lepore; B. C. Smith (eds.), *The Oxford Handbook of Philosophy of Language*, p.250–266. Oxford: Oxford University Press.

Browning, H.; Birch, J. 2022. Animal Sentience. *Philosophy Compass* 17(5):e12822.

Browning, H.; Veit, W. 2020. The Measurement Problem of Consciousness. *Philosophical Topics* 48(1): 85–108.

Browning, H.; Veit, W. 2022. The Sentience Shift in Animal Research. *The New Bioethics* 28: 299–314.

Bugelski, B. R.; Alampay, D. A. 1961. The Role of Frequency in Developing Perceptual Sets. *Canadian Journal of Psychology* 15(4): 205–211.

Burge, T. 1979. Individualism and the Mental. *Midwest Studies in Philosophy* 4(1): 73–122.

Burge, T. 2010. *Origins of Objectivity*. New York: Oxford University Press.

Burge, T. 2022. *Perception: First Form of Mind*. New York: Oxford University Press.

Butlin, P.; Long, R.; Elmoznino, E.; Bengio, Y.; Birch, J. et al. 2023. Consciousness in Artificial Intelligence: Insights from the Science of Consciousness. *Arxiv preprint*.

Carruthers, P. 1992. *The Animals Issue: Moral Theory in Practice*. Cambridge: Cambridge University Press.

Carruthers, P. 2000. *Phenomenal Consciousness: A Naturalistic Theory*. Cambridge: Cambridge University Press.

Carruthers, P. 2005. Why the Question of Animal Consciousness Might Not Matter Very Much. *Philosophical Psychology* 18(1): 83–102.

Carruthers, P. 2017. In Defense of First-Order Representationalism. *Journal of Consciousness Studies* 24(5–6): 74–87.

Carruthers, P. 2018a. The Problem of Animal Consciousness. *Proceedings and Address of the APA* 92: 179–205.

Carruthers, P. 2018b. Comparative Psychology Without Consciousness. *Consciousness and Cognition* 63: 47–60.

Carruthers, P. 2019. *Human and Animals Minds: The Consciousness Questions Laid to Rest*. New York: Oxford University Press.

Carruthers, P. 2020. Stop Caring about Consciousness. *Philosophical Topics* 48(1): 1–20.

Carruthers, P. 2024. *Human Motives: Hedonism, Altruism, and the Science of Affect*. New York: Oxford University Press.

Chalmers, D. J. 1996. *The Conscious Mind: In Search of a Fundamental Theory*. Oxford: Oxford University Press.

Chalmers, D. J. 2018. The Meta-Problem of Consciousness. *Journal of Consciousness Studies* 25(9–10): 6–61.

Chalmers, D. J. 2023. Could a Large Language Model be Conscious? *Arxiv preprint*.

Chemero, A. 2000. Anti-Representationalism and the Dynamical Stance. *Philosophy of Science* 67(4): 625–647.

Chemero, A. 2009. *Radical Embodied Cognitive Science*. Cambridge, MA: MIT Press.

Clark, A. 2015. *Surfing Uncertainty: Prediction, Action, and the Embodied Mind*. Oxford: Oxford University Press.

Clark, A. 2023. *The Experience Machine: How Our Minds Predict and Shape Reality*. New York: Pantheon Books.

Dawkins, M. S. 2021. *The Science of Animal Welfare: Understanding What Animals Want*. New York: Oxford University Press.

Dawkins, R. 1986. *The Blind Watchmaker: Why the Evidence of Evolution Reveals a Universe Without Design*. New York: Norton and Company.

Dehaene, S.; Kerszberg, M.; Changeux, J.-P. 1998. A Neuronal Model of a Global Workspace in Effortful Cognitive Tasks. *Proceedings of the National Academy of Sciences of the United States of America* 95(24): 14529–14534.

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

Dennett, D. C. 1995a. *Darwin's Dangerous Idea: Evolution and the Meanings of Life*. New York: Simon & Schuster.

Dennett, D. C. 1995b. Animal Consciousness: What Matters and Why. *Social Research* 62(3): 691–710.

Devitt, M. 2021. Three Mistakes About Semantic Intentions. In: F. Macagno; A. Capone (Eds.), *Inquiries in Philosophical Pragmatics: Theoretical Developments*, p.9–20. Cham: Springer.

Díaz, R. 2021. Do People Think Consciousness Poses a Hard Problem?: Empirical Evidence on the Meta-Problem of Consciousness. *Journal of Consciousness Studies* 28(3–4): 55–75.

Dung, L. 2022. Does Illusionism Imply Skepticism of Animal Consciousness? *Synthese* 200(3): 1–19.

Edelman, D. B.; Seth, A. K. 2009. Animal Consciousness: A Synthetic Approach. *Trends in Neuroscience* 32(9): 476–484.

Elgendi, M.; Parmod, K.; Barbic, S., Howard, N.; Abbott, D.; Chichoki, A. 2018. Subliminal Priming – State of Art and Future Perspectives. *Behavioral Sciences* 8(6): 54.

Favela, L. H. 2024. *The Ecological Brain: Unifying the Sciences of Brain, Body, and Environment*. New York: Routledge.

Favela, L. H.; Machery, E. 2023. Investigating the Concept of Representation in the Neural and Psychological Sciences. *Frontiers in Psychology* 14: 1165622.

Frankish, K. 2016. Illusionism as a Theory of Consciousness. *Journal of Consciousness Studies* 23: 11–39.

Frankish, K. 2023. What is Illusionism? *Klēsis Revue Philosophique* 55.

Futuyma, D. J. 2021. *How Birds Evolve: What Science Reveals about Their Origins, Lives, and Diversity*. Princeton: Princeton University Press.

Futuyma, D. J.; Kirkpatrick, M. 2022. *Evolution*. 5th ed. Oxford: Oxford University Press.

Godfrey-Smith, P. 2016. *Other Minds: The Octopus, the Sea, and the Deep Origins of Consciousness*. New York: Farrar, Straus, and Giroux.

Godfrey-Smith, P. 2020. Gradualism and the Evolution of Experience. *Philosophical Topics* 48(1): 201–220.

Grice, H. P. 1957. Meaning. *The Philosophical Review* 66(3): 377–388.

Henrich, J. 2021. *The WEIRDest People in the World: How the West Became Psychologically Peculiar and Particularly Prosperous*. New York: Picador.

Henrich, J.; Heine, S. J.; Norenzayan, A. 2010. The Weirdest People in the World? *Behavioral and Brain Sciences*, 33(2–3): 61–83.

Hohwy, J. 2013. *The Predictive Mind*. New York: Oxford University Press.

Hutto, D. D.; Myin, E. 2013. *Radicalizing Enactivism: Basic Minds Without Content*. Cambridge, MA: MIT Press.

Hutto, D. D.; Myin, E. 2017. *Evolving Enactivism: Basic Minds Meet Content*. Cambridge, MA: MIT Press.

Jamieson, D. 1998. Science, Knowledge, and Animal Minds. *Proceedings of the Aristotelian Society* 98(1): 79–102.

Lau, H. 2022. *In Consciousness We Trust: The Cognitive Neuroscience of Subjective Experience*. Oxford: Oxford University Press.

Lee, A. Y. 2022. Degrees of Consciousness. *Noûs* 57(3):553–575.

Levine, J. 1983. Materialism and Qualia: The Explanatory Gap. *Pacific Philosophical Quarterly* 64(4): 354–361.

Machery, E.; Mallon, R.; Nichols, S.; Stich, S. P. 2004. Semantics, Cross-Cultural Style. *Cognition* 92(3): B1–B12.

Mallon, R.; Machery, E.; Nichols, S.; Stich, S. 2009. Against Arguments from Reference. *Philosophy and Phenomenological Research* 79(2): 332–356.

Mashour, G. A.; Roelfsema, P.; Changeux, J-P.; Dehaene, S. 2020. Conscious Processing and the Global Neuronal Workspace Hypothesis. *Neuron* 105(5): 776–798.

Melloni, L.; Mudrik, L.; Pitts, M.; Bendtz, K.; Ferrante, O.; Gorska, U.; Hirschhorn, R.; Khalaf, A.; Kozma, C.; Lepauvre, A.; Ling, L.; Mazumder, D.; Richter, D.; Zhou, H.; Blumenfeld, H.; Boly, M.; Chalmers, D. J.; Devore, S.; Fallon, F.; de Lange, F. P.; Jensen, O.; Kreiman, G.; Luo, H.; Panagiotaropoulos, T. I.; Dehaene, S.; Koch, C.; Tononi, G. 2023. An Adversarial Collaboration Protocol for Testing Contrasting Predictions of Global Neuronal Workspace and Integrated Information Theory. *PLoS ONE* 18(2): e0268577.

Michel, M. 2021. Human and Animal Minds: The Consciousness Laid to Rest. *The Philosophical Review* 130(4): 619–623.

Mylopoulos, M. 2022. Neurobiological Theories of Consciousness. In: B. D. Young; C. D. Jennings (Eds.), *Mind, Cognition, and Neuroscience: A Philosophical Introduction*, p. 280–293. New York: Routledge.

Nagel, T. 1974. What is it Like to Be a Bat? *Philosophical Review* 83: 435–450.

Neander, K. 1991. Functions as Selected Effects: The Conceptual Analyst's Defense. *Philosophy of Science* 58(2): 168–184.

Neander, K. 2017. *A Mark of the Mental: In Defense of Informational Teleosemantics*. Cambridge, MA: MIT Press.

O'Regan, J. K. 2011. *Why Red Doesn't Sound Like a Bell: Understanding the Feel of Consciousness*. New York: Oxford University Press.

O'Regan, J. K.; Noë, A. 2001. A Sensorimotor Account of Vision and Visual Consciousness. *Behavioral and Brain Sciences* 24(5): 939–973.

Papale, P.; Wang, F.; Morgan, A. T.; Chen, X.; Gilhuis, A.; Petro, L. S.; Muckli, L.; Roelfsema, P. R.; Self, M. W. 2023. The Representation of Occluded Image Regions in Area V1 of Monkeys and Humans. *Current Biology* 33(18): 3865–3871.

Papineau, D. 2002. *Thinking About Consciousness*. Oxford: Oxford University Press.

Papineau, D. 2021. *The Metaphysics of Sensory Experience*. Oxford: Oxford University Press.

Papineau, D. (forthcoming) Consciousness is not the Key to Moral Standing. In: G. Lee; A. Pautz (eds.) *The Importance of Being Conscious*. Oxford: Oxford University Press.

Prinz, J. 2012. *The Conscious Brain: How Attention Engenders Experience*. New York: Oxford University Press.

Prinz, J. 2018. Attention, Working Memory, and Animal Consciousness. In: K. Andrews; J. Beck (eds.), *The Routledge Handbook of Philosophy of Animal Minds*, p.185–195. New York: Routledge.

Reber, A. S.; Baluška, F.; Miller Jr. W. B. 2023. *The Sentient Cell: The Cellular Foundations of Consciousness*. New York: Oxford University Press.

Rosenthal, D. M. 1993. State Consciousness and Transitive Consciousness. *Consciousness and Cognition* 2: 355–363.

Rosenthal, D. M. 2005. *Consciousness and Mind*. Oxford: Oxford University Press.

Schiffer, S. R. 1972. *Meaning*. Oxford: Oxford University Press.

Schiffer, S. R. 1982. Intention-Based Semantics. *Notre Dame Journal of Formal Logic* 23(2): 119–156.

Schnell, A. K.; Amodio, P.; Boeckle, M.; Clayton, N. S. 2021. How Intelligent is a Cephalopod? Lessons from Comparative Cognition. *Biological Reviews* 96(1): 162–178.

Schulte, P. 2023. Where Mind Begins: Radical Gradualism about Representation. In: S. F. Souza Filho (ed.), *1° Conferência Virtual de Filosofia da Mente do Recife*.

Schwitzgebel, E. 2023. Borderline Consciousness, When It's Neither Determinately True nor Determinately False that Experience is Present. *Philosophical Studies* 180: 3415–3439.

Searle, J. 1994. Animal Minds. *Midwest Studies in Philosophy* 19(1): 206–219.

Searle, J. 1997. *The Mystery of Consciousness*. New York: New York Review of Books.

Segundo-Ortin, M.; Paco, C. 2023. Plant Sentience? Between Romanticism and Denial: Science. *Animal Sentience* 33(1): 1–32.

Seth, A. K.; Dienes, Z.; Cleeremans, A.; Overgaard, M.; Pessoa, L. 2008. Measuring Consciousness: Relating Behavioural and Neurophysiological Approaches. *Trends in Cognitive Sciences* 12(8): 314–321.

Seth, A. K.; Bayne, T. 2022. Theories of Consciousness. *Nature Reviews Neuroscience* 23: 439–452.

Shea, N. 2018. *Representation in Cognitive Science*. Oxford: Oxford University Press.

Shevlin, H. 2021. Non-Human Consciousness and the Specificity Problem: A Modest Theoretical Proposal. *Mind and Language* 36(2): 297–314.

Simon, J. A. 2017. Vagueness and Zombies: Why 'Phenomenally Conscious' has No Borderline Cases. *Philosophical Studies* 174: 2105–2123.
Souza Filho, S. F. 2022. Can Constancy Mechanisms Draw the Limits of Intentionality? *Disputatio* 14(65): 133–156.
Sytsma, J.; Fischer, E. 2023. 'Experience', Ordinary and Philosophical: A Corpus Study. *Synthese* 201: 210.
Sytsma, J.; Machery, E. 2010. Two Conceptions of Subjective Experience. *Philosophical Studies* 151: 299–327.
Sytsma, J.; Ozdemir, E. 2019. No Problem: Evidence that the Concept of Phenomenal Consciousness is Not Widespread. *Journal of Consciousness Studies* 26(9–10): 241–256.
Tomasello, M. 2023. Social Cognition and Metacognition in Great Apes: A Theory. *Animal Cognition* 26: 25–35.
Tononi, G.; Koch, C. 2015. Consciousness: Here, There and Everywhere? *Philosophical Transactions of the Royal Society B: Biological Sciences* 370(1668).
Tye, M. 2017. *Tense Bees and Shell-Shocked Crabs: Are Animals Conscious?*. New York: Oxford University Press.
Tye, M. 2021. *Vagueness and the Evolution of Consciousness: Through the Looking Glass*. New York: Oxford University Press.
Tye, M. 2024. How Can We Tell if a Machine is Conscious? *Inquiry*: 1–19.
de Waal, F. 2016. *Are We Smart Enough to Know How Animals Are?*. New York: W. W. Norton & Company.
de Weerd, C. R. 2024. A Credence-Based Theory-Heavy Approach to Non-Human Consciousness. *Synthese* 203: 171.
Wynne, C. D. L.; Udell, M. A. R. 2021. *Animal Cognition: Evolution, Behavior and Cognition*. 3rd ed. London: Red Globe Press.

## Notes

[1]Henceforth 'animals'.

[2]I will not address issues of consciousness in other biological kingdoms, such as plants, or in artificial intelligence (AI). For a discussion of plant consciousness, see Segundo-Ortin & Calvo 2023 and the responses to their paper. For a discussion of AI consciousness, see Chalmers 2023; Butlin et al. 2023; and Tye 2024.

[3]Seth et al. (2008) had already treated the measurement problem of consciousness, but only in a broad sense.

[4]As should have been clear, the direct-perception approach appears to transfer certain statements from naïve realism to the context of consciousness attribution.

[5]More recently, Andrews (2024) rejected the epistemic approach in favor of the theory approach.

[6]The term 'Dawkins' in the name of the challenge refers to Marian Stamp Dawkins (2021), a biologist who proposed an argument similar to Carruthers', but one directed at the epistemic approach as a whole.

[7]For more details of Carruthers' arguments, see Carruthers (2019, Chapter 3).

[8]In his earlier works (Carruthers 1992, 2005), although for reasons different from those in his more recent texts, Carruthers also maintained a skeptical stance on animal consciousness.

[9]Other philosophers also share the insight that, beyond human cases, there is no fact of the matter regarding the distribution of phenomenal consciousness (see Papineau 2002; (forthcoming)). In this article, I will restrict myself to Carruthers' position.

[10]Not every gradualist position maintains the conceivability of conscious borderline cases (see Birch 2021). A gradualist could argue that, although consciousness is a phenomenon that admits degrees, there would exist a point at which a non-conscious mental state becomes conscious. That is, upon crossing a non-discrete threshold $P$, a mental state $M_1$ would become, at time $T$, a minimally conscious state $M_2$. For this moderate gradualist, it would be inconceivable for there to be a borderline case of consciousness, insofar as we have a strong intuition that an animal is either at least minimally conscious or not — even if it is acknowledged that the appearance of consciousness in its brain admits degrees. It seems that this gradualist seeks to reconcile the all-or-nothing intuition with the metaphysical gradualist thesis of consciousness. I consider this possibility of reconciliation unsound. The inability to conceive of borderline cases may merely be a matter of how our cognition and perception have been evolutionarily shaped to relate to a world that, for the purposes of survival and reproduction, appears to us as discrete. Cf. Schwitzgebel (2023).

[11]Papineau (2021) and Simon (2017) also seem to share this intuition.

[12]Interestingly, as Dung (2022) suggests, commitment to illusionism does not imply nihilism toward animal consciousness, as advocated by Carruthers. Illusionism, broadly speaking, holds that phenomenal properties do not exist. Our 'qualitative' experience would be nothing more than an *illusion* caused, for instance, by introspective mechanisms that erroneously regard physical properties as being constituted by phenomenal properties. Frankish (2016) refers to these physical properties as '*quasi-phenomenal properties*', where "[a] quasi-phenomenal property is a non-phenomenal physical property (perhaps a complex, gerrymandered one) that introspection typically misrepresents as phenomenal" (Frankish, 2016, p.18). What quasi-phenomenal properties exactly are, according to Frankish (2023, p.6), is an empirical question. They may be, for example, "aspects of sensory processing, features of attentional control, and the wider cognitive effects of sensory information" (*ibid*.). Since the distribution question concerns which organisms instantiate phenomenal properties, the illusionist can shift the debate to which animals possess quasi-phenomenal properties. In this way, the illusionist could address the distribution question of quasi-phenomenal properties. The ultimate issue would then be to investigate which organisms possess introspective mechanisms that misrepresent non-phenomenal physical properties as phenomenal properties, just as, *mutatis mutandis*, occurs with humans. As I see it, the illusionist distribution question could be divided into two problems. The first can be called the *distribution question of deceptive introspection*. This problem would investigate which organisms possess introspective mechanisms capable of misrepresenting physical properties as phenomenal properties. The second can be called the *distribution question of quasi-consciousness*. This problem would aim to investigate which organisms instantiate quasi-phenomenal properties that could be targeted by deceptive introspective mechanisms.

[13]I will not address GWT in detail. For this, see Carruthers (2019); Dehaene (2014); and Mashour et al. (2020).

[14]The mind-body problem is the problem of understanding the relationship between mental states, such as belief, consciousness, and thought, and physical processes in the body (Carruthers 2019, p. 117-118).

[15]Here, I am glossing over many details and issues that arise in this process regarding conceptual and/or non-conceptual content. For example, one thing that perceptual systems must resolve is the 'underdetermination problem', which can be summarized as the following question: "How are perceptual states that represent the environment formed, given that a perceptual system can only access proximal stimulation, and given that this access — together with all the non-psychological laws of physics, chemistry, and biology — determines neither the environmental entities that are represented nor the perceptual states that represent them?" (Burge, 2022, p.51-52). Different authors have proposed distinct approaches to how the perceptual system deals with this. For instance, Burge (2010, 2022) suggests the constancy mechanisms, and Hohwy (2013) and Clark (2015) Bayesian inference as a solution to this problem. However, we can avoid these issues for the purpose of the example presented in the paragraph.

[16]See Carruthers (2018b, p.47) for the presentation of the surprising conditional in his own words.

[17]In $C_n$, '*C*' refers to the phenomenal concept, and the subscript '*n*' denotes the specific token of *C* at a time *t*.

[18]Borg (2008) presents B-style IBS in a slightly different manner. She claims that B-style concerns the determination of an utterance's semantic content. For our purposes, I prefer to interpret B-style in a more specific way — as determining the reference of the utterance — because, as we will see, this aligns closely with how Carruthers embraces IBS regarding phenomenal concepts. Michael Devitt (2021) also shares this interpretation regarding B-style IBS.

[19]Although he is not an eliminativist about qualitative experience, he reduces phenomenal consciousness to access consciousness (Carruthers 2019, p.13).

[20]Thanks to an anonymous reviewer for pressing me to address this point.

[21]Interestingly, on the same page, Carruthers (2019, p. 125) states that, following Kripke, intentions fix the reference, and in the subsequent paragraph he states that it is the *dispositions-to-judge* that fix the reference.

[22]*Quino* is the pseudonym of Mafalda's creator.

[23]To be clear, I am not advocating for a semantic theory that relies on sociological research into people's understanding of concepts like '*p*' or '*q*', and what the cognitive processes responsible for it are. Rather, I am considering that, given Carruthers' strongly naturalistic stance on consciousness, he must provide empirical evidence to justify why he assumes IBS over other naturalistic semantic theories.

[24]For a comprehensive account of the evolutionary history of birds, see Futuyma (2021).

[25]Some authors, such as Peter Schulte (2023), argue that the teleosemantics account of representation should be interpreted in a gradualist manner.

## Acknowledgments

This article is based on a chapter of my master's dissertation. I would like to thank Roberto Horácio, Rodrigo Gouvea, Marco Aurélio, and Sérgio Farias for their helpful comments on an earlier version of this paper. I would also like to thank the anonymous referees for several suggestions to improve the article. This research was funded by a scholarship from CAPES.