revista CONTROVÉRSIA
ISSN 1808-5253
UNISINOS

THE LIMITS OF COMMON SENSE: A NATURALISTIC CRITIQUE OF THE COMMON SENSE IN THE PHILOSOPHY OF CONSCIOUSNESS

DOI: https://doi.org/10.4013/con.2025.213.15

Victor Machado Barcellos
Doutorando em Filosofia pela Universidade Federal do Rio de Janeiro. Financiamento: CAPES.
victorbarcellos1995@gmail.com
https://orcid.org/0000-0003-2269-4923

ABSTRACT:

This paper offers a methodological critique of the material use of philosophical thought experiments and the common sense as evidence in the philosophy of consciousness. As such, I will use Farkas' (2013) paper as a case study. I argue that these kinds of evidence, when used to support metaphysical claims about the nature of conscious states, rely on epistemically unstable judgments that lack sufficient empirical grounding. Drawing on recent findings from experimental philosophy and the social psychology, I propose a moderate restrictionist position, i.e., while not rejecting the method of cases wholesale, we ought to suspend judgment on conclusions drawn fromremote scenarios, particularly in domains where intuitive responses have proven highly variable and culturally contingent. Moreover, I introduce a distinction between objective and hypothetical common sense, arguing that Farkas's appeal to commonsense linguistic practice commits an error. The case study serves to illustrate broader concerns about the evidentiary role of *a priori* intuition and common sense in theorizing about consciousness. I conclude that methodological caution is warranted in the use of material thought experiments and common sense and that future work should prioritize empirically constrained approaches.

KEYWORDS:

Common sense. Philosophy of consciousness. Naturalism. Experimental philosophy.

OS LIMITES DO SENSO COMUM: UMA CRÍTICA NATURALISTA AO SENSO COMUM NA FILOSOFIA DA CONSCIÊNCIA

RESUMO:

Este artigo oferece uma crítica metodológica ao uso material de experimentos de pensamento e do senso comum como evidência na filosofia da consciência. Para tanto, utilizarei o artigo de Farkas (2013) como estudo de caso. Argumento que esses tipos de evidências, quando usados para sustentar teses metafísicas sobre a natureza dos estados conscientes, baseiam-se em julgamentos epistemicamente instáveis e que carecem de fundamentação empírica. Com base em descobertas recentes da filosofia experimental e da psicologia social, proponho uma posição restricionista moderada, isto é, embora não rejeitemos o método dos casos de forma cabal, devemos suspender o juízo sobre conclusões derivadas de cenários remotos, particularmente em domínios nos quais as respostas intuitivas têm se mostrado altamente variáveis e culturalmente contingentes. Além disso, introduzo uma distinção entre senso comum objetivo e senso comum hipotético, argumentando que o apelo de Farkas à prática linguística do senso comum incorre em um erro. O estudo de caso serve para ilustrar preocupações mais amplas sobre o papel evidencial de intuição *a priori* e do senso comum na teorização sobre a consciência. Concluo que é necessária cautela metodológica no uso de experimentos de pensamento materiais e do senso comum e que trabalhos futuros devem priorizar abordagens empiricamente constrangidas.

PALAVRAS-CHAVE:

Senso comum. Filosofia da consciência. Naturalismo. Filosofia experimental.

1 Introduction

The philosophy of consciousness has long relied on bizarre thought experiments to explore the metaphysical nature of conscious states. From zombies (CHALMERS, 1996) and Mary's Room (JACKSON, 1982) to cosmic brains (PAPINEAU, 2021), philosophers routinely construct counterfactual scenarios designed to elicit intuitive judgments about the nature and/or explanatory limits of phenomenal experience. These scenarios, emblematic of what Machery (2017) calls the "material use of the method of cases", have traditionally played a central evidentiary role in arguments for and against physicalism, representationalism, and other metaphysical positions about consciousness.

However, this methodology is increasingly under scrutiny, particularly in light of empirical findings from experimental philosophy and cognitive science. Early studies in experimental philosophy suggested that intuitive responses to philosophical cases vary significantly across cultural, social, and contextual dimensions, raising doubts about their reliability and generalizability (MACHERY*et al*., 2004; WEINBERG*et al*., 2001; SYTSMA, 2014; BUCKWALTER; STICH, 2014). Later research, however, has complicated this picture. Large-scale and replication studies indicate that, in many domains, intuitions prove more invariant across demographic groups than initially assumed (KNOBE, 2019). Taken together, the evidence does not support a simple variability narrative, but rather a mixed pattern, viz., while there are clear cases of demographic divergence, there is also an equally important body of evidence pointing to stability.

In place of this methodology, I defend a form of moderate restrictionism: a domain-specific epistemic policy that urges us to suspend judgment on metaphysical conclusions drawn from radically counterfactual cases in high-variance domains like consciousness. My position does not entail a wholesale rejection of thought experiments, nor a general skepticism toward common sense and/or intuitions. Instead, it advances a naturalistically informed stance that prioritizes empirical tractability, psychological realism, and methodological conservatism, especially when theorizing about mental phenomena that are themselves experientially opaque and conceptually elusive.

It should be noted that some philosophers, such as Williamson (2021), Bengson (2014), and Chudnoff (2013), have advanced sophisticated defenses of the evidential legitimacy of intuitions for the evaluation of philosophical theses in general. These accounts typically treat intuitions as a form of intellectual seeming, or as defeasible but nonetheless *prima facie* justification, thereby integrating them into a broader epistemology conception of knowledge. To engage directly with this methodological debate would require adjudicating questions concerning the nature of intuitions, their epistemic status, and their role in the architecture of justification. However, this is something beyond the scope of the present paper. My concern here is more restricted and domain-specific, viz., the role of judgments and commonsense in the philosophy of consciousness. Even if one grants that intuitions may possess evidential weight in certain philosophical domains, it does not follow, I believe, that they maintain such legitimacy in the philosophy of consciousness. As we shall see, Rolla's (2021) pragmatic view helps to clarify why: the reliability of conceptual competences is bounded by the historical-practical contexts in which they were acquired by the subject, and their extension to radically unfamiliar scenarios, precisely the kind that populate consciousness studies, cannot be methodologically vindicated.

That said, to motivate and develop this critique, the paper proceeds as follows. Section 2 clarifies the typology of philosophical cases, drawing on Machery's framework to distinguish between illustrative, formal, and material uses. Section 3 reconstructs Farkas's theory of exteroceptive pain and her reliance on both thought experiments and appeals to common sense. Section 4 critiques the evidentiary force of the material cases Farkas deploys, particularly with respect to their lack of empirical constraint and susceptibility to intuitive variance. Section 5 challenges her appeal to common sense by distinguishing between *objective* and *hypothetical* common sense, and by drawing on empirical findings from experimental philosophy of perception to show that folk intuitions are, in fact, more relational than her argument presupposes. Finally, the paper concludes by proposing a methodological constraint: philosophers working on consciousness should subject their intuitive and common-sense judgments to

empirical triangulation, and in the absence of such triangulation, adopt a cautious stance toward metaphysical theorizing based on armchair reflection alone.

2 The Method of Cases

The deployment of cases for diverse epistemic and argumentative purposes is a pervasive methodological strategy in analytic philosophy, and the philosophy of consciousness is no exception. But what exactly constitutes a "case"? As a first approximation, Machery (2017, p. 17) characterizes cases as “descriptions of actual or hypothetical situations”, noting that philosophical cases are simply those presented or constructed by philosophers for theoretical purposes. Typically, such cases are designed to elicit intuitive judgments from the reader regarding the plausibility, coherence, or implications of the scenario under consideration. A paradigmatic example is Jackson’s (1982)Mary’s Room thought experiment. In this case, we are asked to imagine Mary, a neurotypical human subject, raised in a strictly black-and-white environment but endowed with exhaustive knowledge of the neurophysiological processes underlying color perception. Upon leaving her achromatic confinement and encountering a red object for the first time, the central question arises: does Mary acquire new knowledge, specifically, knowledge of what it is like to see red? The thought experiment aims to evoke a judgment about the cognitive significance of phenomenal experience, thereby challenging the completeness of physicalist theories of consciousness. Although philosophical cases provoke an attitude, they remain neutral regarding what holds in the scenario described.

Philosophical cases serve a variety of functions within philosophical inquiry. According to Machery’s taxonomy (2017, Ch. 1), these include provocative, illustrative, formal, material, and exploratory uses. Let’s briefly see each of them.

The provocative use aims to catalyze philosophical reflection by eliciting intuitive judgments that give rise to conceptual puzzles or theoretical tensions. It functions as a heuristic entry point, often remaining neutral with respect to any particular philosophical thesis. The illustrative use, by contrast, seeks to render arguments, theories, or analyses more accessible by embedding them in imaginative scenarios. Both of these uses are considered philosophically innocuous in Machery’s framework. The reason, according to Machery, is that they neither presuppose nor entail the truth of the philosophical claims under discussion.

In contrast, formal and material uses are philosophically substantive. When employed formally, a case is intended to illuminate the semantic content or logical structure of a target concept. It may expose the necessary and/or sufficient conditions for the application of that concept or reveal conceptual tensions

latent in its usage. Material uses, on the other hand, are directed not toward concepts but toward metaphysical features of the phenomenon itself. For instance, the Mary thought experiment in the material sense aims to test the adequacy of physicalism by probing whether phenomenal consciousness can be exhaustively explained in physical third-person terms.

This paper focuses specifically on the material deployment of philosophical cases. It is important to emphasize that the same scenario may be repurposed for different epistemic functions depending on its discursive context (GIRALDO, 2025). For example, while Machery (2017, p. 15) classifies Davidson's (1987) Swampman scenario as merely illustrative within the original context of Davidson's argument, aimed at discussing issues of meaning, it has subsequently been appropriated in debates over teleosemantics as a canonical counterexample that purports to undermine the theory's core assumptions (see Papineau, 2022). In the Swampman thought experiment, Davidson asks us to imagine that he is struck by lightning in a swamp and instantaneously replaced by a molecule-for-molecule duplicate (a Swampman) constructed out of swamp matter. Although Swampman is physically identical to Davidson and behaves indistinguishably, it lacks Davidson's causal history, i.e., it has no learning history and no evolutionary background. Teleosemantic theories of mental content hold, roughly speaking, that mental representation depends on proper functions established via natural selection and/or learning history. In short, a mental state represents R_i in virtue of having been selected *for* its capacity to indicate or respond to R_i in a functionally advantageous way (MILLIKAN, 1984). However, since Swampman lacks any such etiological history, it follows from teleosemantics that its internal states lack representational content, even though, intuitively, Swampman sees, thinks, and reasons just like Davidson. The material use of the Swampman case thus aims to generate a *reductio ad absurdum* of the teleosemantic framework, i.e., if the theory cannot attribute content to a being that is functionally and physically indistinguishable from a human, then the theory is arguably missing something fundamental about the nature of mental content.

The aim of this paper is not to reject the legitimacy of the method of cases in philosophy *tout court*, nor to dismiss the usefulness of various case-driven strategies within the philosophy of consciousness. Indeed, I take it that the illustrative use of thought experiments, such as employing scenarios to clarify conceptual distinctions or illuminate theoretical commitments, can be epistemically productive. My concern here is more restricted: it targets the material deployment of the method of cases within consciousness studies. In particular, I am interested in its construction as a remote scenario, where such cases are typically designed to support or undermine substantive metaphysical theses. Here I will rely on Rolla's (2021) pragmatic view to constrain the evidential role of intuitions in debates on consciousness. Although Rolla develops this view in the context of epistemology, I shall, *mutatis*

mutandis, transpose it to the philosophy of consciousness. Rolla argues that the reliability of intuitions is limited by the historical and practical conditions that shape concept use. On this account, the deployment of a concept is inseparable from the practices in which it was historically refined, practices that originated in concrete classificatory and explanatory difficulties. Accordingly, an intuitive judgment concerning the application of a concept is reliable only to the extent that it preserves continuity with the contexts of practice that originally conferred competence in its use. This stance is grounded in two independently plausible theses: (a) mastery of a skill results from its practice, and (b) the success of a skill depends on the context in which it is exercised. In other words, if an agent *S* in a context *C* at time t_1 successfully masters a skill *H* through practice in *C*, then *S*'s reliable deployment of *H* at a later time t_2 is constrained to contexts *C′* that are sufficiently continuous with *C*. In contexts where *C′* is radically different from *C*, *S* cannot be expected to exercise *H* reliably at t_2 (ROLLA, 2021, p. 24).Just as the ability to swim, while effective across a wide range of aquatic contexts fails in radically different environments such as arid deserts, where the very medium that enables the exercise of the skill is absent. In an analogous way, the competence underlying concept use cannot be expected to transfer to radically unfamiliar scenarios, where the historical and practical conditions that originally shaped the concept's application no longer obtain. In the case of consciousness, this implies that intuitive judgments elicited by remote scenarios cannot be trusted, since they are detached from the historical-practical contexts that gave rise to our competence with the concept.

This diagnosis helps illuminate why certain canonical thought experiments in the philosophy of consciousness are methodologically problematic. For example, Papineau (2021) appeals to the cosmic brain case, a spontaneous physical duplicate of a neurotypical human brain arising *ex nihilo* in space that is purported to undergo conscious states, in defense of an intrinsicalist theory of consciousness. Likewise, Chalmers (1996) famously invokes the philosophical zombie, a molecule-for-molecule and functional duplicate of a conscious agent that lacks phenomenal experience, to argue against physicalist accounts of consciousness. These cases exemplify the reliance on remote scenarios whose intuitive appeal is precisely what Rolla's pragmatic stance renders suspect. Importantly, however, Rolla (2021, p. 25) does not claim that intuitions are *tout court* unreliable or that they should be eliminated from philosophical practice. Rather, intuitions remain methodologically appropriate when applied to ordinary situations continuous with the contexts that originally shaped conceptual competence. What is problematic is their extension to radically remote cases, where that continuity no longer obtains. My aim, therefore, is not to deny the

heuristic value of thought experiments altogether, but to articulate a principled methodological constraint on their material deployment in consciousness studies.[1]

In the next section, I turn to a concrete case study to illustrate how this methodological constraint should operate in practice. By examining a specific instance, we can see more clearly how Rolla's pragmatic view, when transposed to the philosophy of consciousness, guides the evaluation of thought experiments and delimits both the evidential role of intuitions and appeals to common sense.

3 Farkas' Argument: Pain and Exteroception

Katalin Farkas's (2013) article will serve as a case study for evaluating methodological practices in the philosophy of consciousness. I contend that her work exemplifies two argumentative strategies that, while widespread in the literature, are epistemically problematic, namely, (i) the material use of the method of cases constructed as remote scenarios (ii) the appeal to what is purportedly commonsense as evidence for substantive philosophical conclusions. The selection of this article is not meant as a critique of its metaphysical conclusion, but rather as an illustrative example of how debates in the field can be distorted by these methodological tendencies.

One of Farkas's central aims is to argue that there exists a subclass of pain experiences that are constitutively exteroceptive perceptual states (2013, p. 102). She characterizes perceptual intentionality as a feature of conscious states that meet the following two conditions: (i) the state occurs in a sensory modality, and (ii) it presents an experience-independent world, populated by experience-independent objects and properties (2013, p. 100). In other words, for a conscious mental state to possess perceptual intentionality, it must *seem* to be directed at a mind-independent external reality. This reflects a commitment to what Pautz (2021) has labeled Essential External Directedness, the thesis that perceptual experience is essentially characterized by an apparent orientation toward entities and properties external to the subject. On this view, perceptual states are not defined merely by their sensory character, but by their intentional structure, that is, by the way they purport to represent a world that transcends the subject's experiential perspective. Hallucinatory experiences, for example, would still count as perceptual (in this sense) because they present themselves *as if* directed at something in the external world, even if no such object is present. According to Farkas, my visual experience of the computer in front of me qualifies as a perceptual experience because it arises through the visual modality and seems to present a physical object (a black computer with a determinate shape and spatial location) that exists independently

[1] For in-depth discussions on what exactly the case method is and how it is employed in philosophy, see Giraldo (2025) and Machery (2017).

of my perceptual engagement with it. Further, Farkas (2013, p. 102) stipulates that a perceptual experience is exteroceptive just in case the perceptual object is an external cause of the perceptual state. This causal requirement distinguishes exteroception from interoception or proprioception, which concern internal bodily or kinesthetic states. Her goal is to show that some pains, typically regarded as paradigmatically interoceptive, may actually satisfy these criteria for exteroceptive perception.

Farkas distinguishes between two types of pain experiences. The first, called *Pain1*, is the experience characterized by its own experiential quality of being painful. Pain1 is, in this sense, individuated by its intrinsic (i.e., non-relational) property of being a painful experience for the subject undergoing it. For example, when I cut my hand with a knife, my conscious experience is characterized by having a negative hedonic property[2], viz., being painful. This mental state, *prima facie*, does not involve intentional directedness toward the object (e.g., the knife) that caused the nociceptive stimulation.[3] According to Farkas, the experiential quality Pain1 is a property "that can be exemplified *only* if a conscious creature has the appropriate kind of experience" (2013, p. 104, emphasis added).In other words, Pain1involves a hedonic quality that is *experience-dependent*, and thus essentially tied to the presence of a conscious subject.

Pain2, by contrast, would be a case of exteroceptive perception, in which the perceptual experience tracks a negative hedonic property that is experience-independent. On this view, Pain2 would be a property that exists independently of a subject who experiences it. Just as one might perceive the shape or color of a pencil as objective features of the object, Pain2 purports to track an external, mind-independent painfulness. In adopting this view, Farkas appears to endorse a form of realism and anti-relationalism about certain hedonic properties, namely, that some unpleasantness is not merely a subjective projection, but a feature of the world. Thus, these properties are intrinsic to the object that instantiates them.[4] To make this position precise, we can define what I will call the Anti-Relationalist

[2]Roughly speaking, a *hedonic property* is a phenomenal property characterized by either positive or negative *valence*. Paradigmatic examples include pleasure (positive valence) and pain (negative valence). The nature of valence is a matter of ongoing philosophical debate, with at least four major theoretical approaches: intrinsic-feeling theories, which treat valence as a primitive phenomenal quality (MOORE, 1903); representationalist theories, which analyze valence in terms of the content of evaluative representations (CARRUTHERS, 2024); imperativist theories, which hold that valenced states issue motivational commands (KLEIN, 2015); and enactivist theories, which ground valence in embodied affective dynamics (COLOMBETTI, 2014). For present purposes, however, we need not engage with the details of this debate.

[3] This is contentious if we take *being directed at* to mean *representing a state of affairs*. There are several representationalist theories of pain according to which pain represents, e.g., the injury as being painful. See Cutter & Tye (2011).

[4] As a rule, anti-relationalist positions reduce pain to physical properties of the body, such as actual or potential bodily damage, which are represented as negative by the subject's pain system (TYE, 2000). Farkas goes even further by claiming that pain properties can be instantiated by entities such as a corridor or a leaf. Pautz (2014) argues that anti-relationalism conflicts with some strong intuitions, such as the idea that pain cannot exist without someone experiencing it (the *esse est percipi* intuition). Moreover, he finds it odd that the same unpleasant quality you experience when cutting your hand could be instantiated in a

Theory of Pain-Property (ARTPP). According to ARTPP, a property *P* is a hedonic property of an object *O* if and only if: (i) The existence of *P* and *O* does not depend on any relation or set of relations *R* involving a conscious experience E_i instantiated by a subject of experience *S*; and (ii) *P* is not a merely dispositional property, i.e., it is not merely the potential to cause a canonical experience E_i in a typical subject *S* under standard conditions. In short, on this anti-relationalist account, there are intrinsically negative properties in the external world that are perceptible, and Pain2 is the kind of perceptual state that discloses them. This is a strong metaphysical thesis with substantive implications for the ontology of affective states and the intentional structure of conscious experience.[5]

As previously stated, Farkas supports the ARTPP through two main argumentative strategies. The first involves the material use of the method of cases. In particular, she appeals to two thought experiments that, according to her, exemplify cases in which pain is perceptually presented as an exteroceptive property, that is, as an experience-independent hedonic feature instantiated by external particulars in the world.

The first case is drawn from §312 of Wittgenstein's *Philosophical Investigations*. Wittgenstein invites us to imagine a hypothetical world in which "[t]he surfaces of the things around us (stones, plants, etc.) have patches and regions which *cause pain in our skin* when we touch them." He adds: "In this case, we'd speak of pain-patches on the leaf of a particular plant, just as we speak of red patches." (2009 [1953], §312, p. 110[e], emphasis added). On this interpretation, pain is attributed in the same epistemic and semantic register as perceptible qualities, suggesting that it is treated as an objective property of external surfaces. The second case, attributed to A. J. Ayer, involves the following scenario: Suppose there is a large building containing a complex network of corridors.[6] One particular corridor systematically and reliably induces an intense headache as soon as one enters it. Each time the subject walks into the corridor, the headache arises immediately, and each time the subject exits the corridor, the headache ceases just as promptly. After several such encounters, the subject begins to describe the space as having

corpse (PAUTZ, 2014, pp. 239–240). As I have made clear, my goal here is not to refute Farkas's anti-relationalism, but rather to highlight the methodological limitations of the arguments that support it. In short, my argument is not metaphysical, but epistemic.

[5] Ultimately,Pain2appears to belong to the class of primary qualities, as formulated by John Locke (1975 [1690]). Locke distinguishes between primary and secondary qualities. Arguably, primary qualities, such as solidity, extension, figure, and motion, are properties that exist in objects independently of any observer. They are objectively measurable and thought to resemble the way the world actually is. In contrast, secondary qualities, like color, taste, or sound, depend on the interaction between the object and a perceiving subject. It is interesting to note that Farkas uses the term "quality" rather than "property". Here, I am using "property" because I consider it a more neutral term than "quality", given that the latter, in the context of philosophy of mind, typically refers to something of a mental nature. However, as Farkas seems to indicate, Pain2 is not a feature of mental states, but rather of objects external to the mind. Thus, in this context, it seems more appropriate to use the term "property".

[6] I say "attributed" because Farkas does not provide a reference indicating where Ayer presents this thought experiment.

a distinctive hedonic profile, for instance, by saying, "*Let's avoid the headachy corridor*". Farkas (2013, p. 102) contends that this is analogous to how we describe corridors as dark, noisy, or smelly. Just as these adjectives describe perceptual properties of external environments, the term "headachy" is taken to describe a negative hedonic quality that is perceived as belonging to the corridor itself.

According to Farkas, both cases exemplify experiences that meet her criteria for perceptual intentionality: they are (i) sensory in mode and (ii) apparently directed at experience-independent properties of external objects or locations. Thus, on her view, these cases support the plausibility of ARTPP.

4 Critique of the Material use of the Method of Cases in the Philosophy of Consciousness

Let us consider Wittgenstein's thought experiment. Farkas interprets §312 of the *Philosophical Investigations* as suggesting that certain surfaces, such as a leaf, could instantiate experience-independent negative hedonic properties, described as "pain patches". According to Farkas, just as tactile perception can present secondary qualities like roughness or warmth, it could likewise present painfulness as a distal property of external objects, directly accessible through exteroceptive sensory modalities. That is, the leaf of the plant does not merely cause pain in us due to some nociceptive stimulation, rather, it instantiates a mind-independent painfulness in the same way it may instantiate color, shape, or texture. In her reading, this supports the thesis that certain pain experiences are constitutively exteroceptive perceptions, presenting intrinsic affective properties in the environment.

However, Wittgenstein's paragraph does not definitively establish whether the relevant property is (i) an intrinsic, mind-independent negative hedonic quality instantiated by the object (as Farkas assumes), or rather (ii) a dispositional property, whose instantiation depends on its causal potential to elicit a nociceptive experience in a perceiver. The passage emphasizes that the patches "*cause* pain", a formulation more consistent with a causal-dispositional or relationalist framework. On this view, painfulness is not a monadic property inhering in the object per se, but a secondary quality, realized through structured relations between the object's physical profile and the affective sensitivities of embodied agents capable of experience pain. In other words, the thought experiment may be read not as positing the perceptibility of Pain2 as an intrinsic, experience-independent property instantiated in the object, but rather as illustrating a scenario in which an agent would judge an object to be painful in virtue of its causal role in producing an unpleasant phenomenal state upon injury. Pain2, so interpreted, is a dispositional property that becomes manifest only when the object interacts with the bodily schema of a sentient organism in a way that results in the distinctive qualitative character of pain. Thus, what the

scenario arguably reveals is not a *sui generis* perceptible property instantiated in external objects, but a contingent disposition to elicit certain types of first-person affective states under specific conditions.This interpretive indeterminacy undermines any direct metaphysical inference from the scenario.In the absence of further conceptual argumentation or empirical triangulation, I suggest that the intuitions evoked by the scenario lacks the requisite inferential determinacy to ground the anti-relationalist conclusion Farkas draws from it.

To be clear, my aim is not to resolve the exegetical question on Wittgenstein's text, a task better left to scholars of his philosophy. My concern, rather, is to question the status of the facts assumed to obtain in the thought experiment. In other words, the facts assumed to obtain in Wittgenstein's thought experiment are themselves unclear, and this uncertainty directly undermines the metaphysical conclusion that Farkas seeks to draw from it.

Nevertheless, one could argue that there is currently no body of empiricalresearch assessing the judgments of laypeople or experts regarding this specific case. If that is the case, so we cannot argue that the intuitions evoked by Wittgenstein's thought experiment is demographic variable (as occurs with other cases). However, given the growing evidence that judgments elicited by other philosophical thought experiments in the philosophy of mind are often variable, culturally biased, and sensitive to superficial framing effects (MACHERY, 2017, Ch. 2; SYTSMA, 2014), it would be methodologically imprudent to grant special epistemic authority to our intuitions in this case.[7] If intuitions elicited by similar cases, e.g., zombies, Mary's room, Swampman, are subject to epistemic noise, then by parity of reasoning, Wittgenstein's pain patch scenario is likely similarly vulnerable. As such, the most epistemically cautious and parsimonious stance is one of suspension of judgment regarding which metaphysical facts hold in the thought experiment, particularly in the absence of corroborating empirical constraints.For instance, Buckwalter and Stich (2014) presented participants with a series of metaphysical scenarios, including brain-in-a-vat, Chinese room, Twin Earth, and Mary's room, and found that gender was a significant predictor of intuitive response patterns. Women were statistically more likely than men to affirm knowledge or agency in skeptical scenarios, suggesting that intuitions about canonical philosophical thought experiments may not be universally shared but are modulated by social and psychological variables. Although subsequent studies have debated the implications of these findings, they nonetheless provide evidence that intuitions in the metaphysics of mind are not epistemically uniform across different

[7] It should be noted, however, that not all studies converge on the fragility of philosophical intuitions. As discussed earlier, more recent work has shown that in a number of domains intuitions exhibit a significant degree of stability across demographic groups. My point here is not to deny such robustness, but to emphasize that in the philosophy of mind (and particularly in the context of remote scenarios) there remains substantial evidence that intuitions can be sensitive to demographic, cultural, and framing factors, which cautions against treating them as universally reliable guides.

groups.Further evidence of intuitive variance comes from Fischer and Sytsma (2021), who investigated the effects of subtle linguistic framing on participant's judgments about philosophical zombies. In one condition, participants were asked whether a "zombie" was capable of experiencing a conscious state such as feeling emotions or if the zombie would be capable of have an inner life. In another condition, the entity was described as a "perfect physical duplicate" of a human being, without using the term "zombie". The participants of the perfect physical duplicate scenario were asked the same questions that were offered in the zombie condition. Despite the scenarios being metaphysically equivalent, participant's responses diverged significantly across conditions. That is, those in the zombie condition were more likely to deny phenomenal consciousness to the duplicate, while those in the perfect physical duplicate condition were more likely to affirm it. This result suggests that even slight shifts in phrasing can elicit different intuitions. This may highlight the fragility and context-dependence of the judgments evoked by thought experiments used in the philosophy of consciousness.

In light of such findings, it would be epistemically reckless to assume that the pain patch scenario is immune to similar forms of noise. If paradigmatic thought experiments concerning consciousness, such as zombies, Mary's Room, and the brain-in-a-vat, reliably produce discordant and theoretically contaminated intuitions, then,*prima facie*, the Wittgensteinian example is unlikely to constitute an epistemic outlier. By parity of reasoning, and in the absence of corroborating empirical constraints, the most methodologically responsible stance toward the metaphysical structure of the imagined case is one of suspension of judgment. Such caution is not a retreat into skepticism, but rather a principled commitment to evidential discipline in the face of underdetermination. Moreover, Rolla's (2021) pragmatic view reinforces this conclusion: since the reliability of intuitions is bounded by the historical-practical contexts in which conceptual competences were originally acquired, their extension to remote scenarios like the pain patch cannot be methodologically vindicated. Indeed, the very concept of "pain" is not historically bounded or contextually practiced in ways that allow for its coherent deployment in corridors, plant patches, or similar contrived environments, making such applications especially vulnerable to breakdowns in reliability. Such caution is not a retreat into skepticism, but rather a principled commitment to evidential discipline in the face of underdetermination.

The position I am defending is not a form of *radical restrictionism* as advocated by Machery (2017), but rather a version of *moderate restrictionism*. Specifically, I argue that, in light of empirical evidence from x-phi of mind and consciousness, as well as relevant findings from social psychology, we ought to adopt a policy of localized epistemic suspension toward a significant subset of philosophical thought experiments, viz., those arising within the domain of the philosophy of consciousness. Unlike

Machery's global critique, which targets the method of cases across the entirety of philosophical practice, my critique is domain-specific and thus avoids the methodological overreach of generalizing empirical results from one philosophical area to others. Because my restrictionism is local rather than global, it escapes the problem of illicit extrapolation. Machery's strategy requires that we treat data drawn from relatively well-studied domains (e.g. epistemology, semantics, moral theory) as representative of philosophical methodology as a whole, including underexplored or methodologically distinct fields such as aesthetics or metametaphysics, areas for which experimental data is sparse or nonexistent. As Machery himself concedes, such extrapolation exposes his argument to epistemic vulnerability, viz., the presumption that domains with minimal empirical scrutiny exhibit the same cognitive biases and inferential unreliability as domains with robust empirical coverage. By contrast, my moderate restrictionism is empirically grounded, methodologically conservative, and scope-limited. It leverages data from the x-phi of mind and consciousness, and thus avoids both overgeneralization and the burden of universal methodological critique. Even in cases where empirical studies of specific thought experiments are lacking, my argument is bolstered by patterned findings across related cases within the same subfield (e.g., zombies, Mary's room, brain-in-a-vat, twin earth, etc.), supporting the conclusion that the use of the material sense of the method of cases in consciousness studies is particularly prone to epistemic unreliability. For this reason, my call for suspending judgment is, I suggest, epistemically justified and methodologically proportionate.

5 Empirical Evidence Against Hypothetical Common Sense

Farkas's second argumentative strategy consists in appealing to common sense linguistic intuitions as a form of evidential support for her metaphysical thesis. Specifically, she appears to justify the claim that there are two ontologically distinct kinds of pain experiences, one that is intrinsically subjective (experience-dependent), and another that is perceptually directed at objective properties (experience-independent), by drawing an analogy with taste perception. According to Farkas, ordinary language users routinely deploy two notions of "taste": one that is phenomenally grounded, referring to the subjective flavor experience of the taster, and another that attributes the property of "tastiness" or "flavor" directly to the object, as if it were an intrinsic feature of the item being tasted. This dual usage in everyday discourse is then marshaled as support for the claim that people implicitly recognize a distinction between subjective and objective sensory qualities, an observation which, in her view, lends

credence to the idea that certain pain experiences may likewise be understood as directed at objective, experience-independent properties of external objects. In her own words:

> If someone has a bad head-cold, *we sometimes say* that she cannot taste anything. But *people occasionally also say* when they have a cold that everything is completely tasteless. There are *two different notions* here: *Taste1 is a feature that can be exemplified only when a conscious creature has the appropriate experience*, whereas *Taste2 is a quality that can be exemplified even in the absence of any such experience*. (FARKAS, 2013, p. 104, emphasis added)

Note that Farkas appears to treat commonsense as justificatory evidence for a substantive metaphysical thesis. However, the appeal to commonsense judgments as epistemically authoritative is a highly contentious practice in contemporary philosophy of perception and, to a lesser extent, in other branches of philosophy (see BLOCK, 2023; BURGE, 2022, Ch. 1; CHEMERO, 2009, Ch. 1). After all, what grounds the assumption that commonsense discourse reliably tracks the metaphysical structure of the world, particularly with respect to the distinction between experience-dependent and experience-independent properties? This is a significant epistemological question that Farkas does not engage. She appears to take for granted that the intuitions or linguistic practices of non-experts offer credible support for positing the existence of objectively instantiated hedonic properties. That said, my critique does not concern the use of commonsense *per se*, either as a legitimate evidentiary source or as a starting point for philosophical theorizing.[8] Rather, my concern is directed at what may be termed "hypothetical commonsense". This refers to a methodological strategy in which philosophers, without the benefit of empirical validation, attribute specific beliefs or intuitions to "most people" based primarily on introspection, anecdote, or sociocultural familiarity. In other words, let *HCS*(*p*) refer to the claim that proposition *p* is endorsed by "common sense". Then *HCS*(*p*) is allowed iff a philosopher *T* attributes *p* to a majority of ordinary agents *A* without independent empirical evidence of *A*'s actual judgments. I call it "hypothetical" because its evidentiary basis often consists of nothing more than the philosopher's own assertions, assertions which may reflect personal socialization, cultural background, or academic milieu, rather than data-driven generalizations. As such, hypothetical commonsense risks functioning as projected folk metaphysics, whose authority remains unsubstantiated.

However, in my view, none of the aforementioned considerations establish that the propositions attributed to "commonsense", when extrapolated from the philosopher's own armchair perspective,

[8] The debate over whether *taste* and *flavour* properties belong intrinsically to the object (e.g., whether the sweetness of wine is a property that exists independently of a subject who experiences it), whether they are emergent properties of a certain arrangement of chemical compounds but not reducible to them, or whether they are relational properties that depend on the subject of experience, is longstanding and controversial. Typically, such discussions are guided either by empirical evidence from psychology and neuroscience or by *a priori* arguments. Since my argument is not metaphysical in nature but epistemic, I will not engage with this debate.

genuinely reflect the inferential or doxastic attitudes held by non-specialists. Appeals to this kind of hypothetical commonsense, that is, to putatively widespread intuitions posited without empirical warrant, seem increasingly methodologically suspect, particularly given the development of dedicated empirical frameworks, such as experimental philosophy, aimed precisely at investigating the content, distribution, and stability of folk judgments regarding philosophically salient cases. What I propose to call *objective commonsense* refers to intersubjectively available judgments elicited from representative populations (or subpopulations), obtained through systematic methodologies such as survey experiments, vignettes, and psychometric analysis. These data-driven approaches stand in sharp contrast to *a priori* projections about what "most people" allegedly believe. Reliance on hypothetical commonsense reverts, in effect, to a pre-empirical epistemology of intuitions, where philosophical authority is grounded in introspective plausibility or cultural familiarity rather than in evidence of any kind. *Mutatis mutandis*, this is the very form of speculative methodology that the naturalistic turn in philosophy, including but not limited to x-phi, sought to displace.[9] Indeed, the continued appeal to hypothetical commonsense in certain metaphysical arguments, as exemplified in Farkas's work, appears methodologically anachronistic. It presupposes, without validation, that a philosopher's first-person epistemic access to their own linguistic and social environment suffices to generalize across communities of speakers. Early work in experimental philosophy did suggest that such generalizations are problematic, since intuitive responses can be modulated by variables such as culture, education, or socio-linguistic context (MACHERY*et al*., 2004; WEINBERG*et al*., 2001; SYSTMA, 2014). More recent studies, however, have revealed a more nuanced picture: in a number of domains, philosophical intuitions exhibit a significant degree of stability across demographic differences (KNOBE, 2019). What follows is that the mere reliance on one's own intuitions as an epistemic proxy for the community at large is methodologically questionable, regardless of whether intuitions prove stable or variable in particular cases. In this respect, Farkas's approach seems to replicate a pre-scientific posture, in which the philosopher is treated as a stand-in for the linguistic practices of an entire community, an assumption that is untenable in light of the evidential standards set by contemporary cognitive and social science.

But are there any pieces of evidence indicating that people in fact judge that there are taste properties that are experience-independent, as Farkas suggests? Let us examine some of them.

In their influential article on the experimental philosophy of perception, Cohen and Nichols (2010) empirically investigated the introspective rejoinder argument in the context of the metaphysics of color.

[9] Unfortunately, the arm chair approach that relies on hypothetical common sense still seems to be popular among philosophers, even in areas of philosophy where we already have well-established and rigorous sciences, as is the case with the philosophy of perception.

Broadly speaking, there are two dominant theoretical views in the philosophy of color: relationalism and anti-relationalism.[10] Relationists maintain that color properties are constitutively relational, i.e., their metaphysical nature is partially determined by relations to perceiving subjects and environmental conditions. For example, the color blue is understood as a property that arises from a systematic causal relation between an object's surface reflectance profile and the sensory-processing architecture of a perceptual system (whether human or otherwise). By contrast, anti-relationists hold that colors are intrinsic, perceiver-independent properties (typically surface reflectance properties) that objects instantiate regardless of whether they are observed. Thus, in a hypothetical observerless world, relationists would deny the existence of color, while anti-relationists would affirm it.

The introspective rejoinder is commonly invoked in support of anti-relationalism. According to this argument, introspective access to the phenomenology of color perception reveals colors as intrinsic properties of external objects(WATKINS;SHECH, 2025, p. 20). For instance, when I perceive the redness of a tomato, my experience ostensibly presents redness not as a relational feature contingent upon my perceptual apparatus, but as a stable, mind-independent feature of the tomato's surface. In this sense, introspection is said to provide direct insight into the metaphysical nature of color (TYE, 2000, pp. 152-153). The claim is that color experience presents color as objective in the same way shape or size is typically presented.

To empirically probe this line of reasoning, Cohen and Nichols (2010) constructed a series of vignettes involving perceptual disagreements between a human and an alien observer. They presented participants with a series of representational variations involving shape, color, taste properties, and the application (or not) of the property *delicious*.[11] In each variation, a scenario was described involving a human and an alien whose perceptual system was entirely different from ours.[12] For example, in one of these scenarios, it was described that Andrew, an alien, and Harry, a human, were both observing a ripe tomato in good lighting at a distance of one meter. Harry reported that the ripe tomato was red, whereas Andrew claimed that it was green. After the scenario was described, participants were asked to select one of the following three responses: (i) The tomato was red, and therefore Harry was right and Andrew was wrong; (ii) The tomato was not red, and therefore Andrew was right and Harry was wrong; or (iii) There is no fact of the matter regarding claims such as "the tomato is red". That is, different people have

[10] Although the philosophical debate between relationalism and anti-relationalism regarding perceptible properties is largely dominated by the ontology of color, it also extends to, e.g., warmth and sound.

[11] The participants were undergraduate students from an introductory logic course ($n = 31$).

[12] As a way to avoid interpreting the result as due to linguistic differences, the researchers presented a prior description in which the aliens had learned and become fluent in English.

different visual experiences when perceiving the same object, and thus, there is no absolute truth to the claim "the tomato is red" (COHEN;NICHOLS, 2010, p. 222). The researchers predicted a higher rate of responses (i) and (ii) if participants held anti-relationist intuitions, and (iii) if they held relationist intuitions. Regarding the responses that interest us, the majority of participants chose (iii) for the property *taste*. As for the property *delicious*, which bears similarities to *taste*, most participants also responded with (iii) (FIGURE 1).

In another experiment, Roberts and Schmidtke (2016) conducted a series of studies aimed at determining whether people have relationalist or anti-relationalist intuitions regarding auditory and taste properties.[13] For our present purposes, I will focus on two of the items analyzed, specifically those related to taste properties. In both experiments, participants were presented with a scenario description followed by a table containing four philosophical statements: metaphysical, ascription, realism, and correspondence. Participants were asked to indicate their level of agreement with each statement on a scale from 0 (Strongly Disagree) to 10 (Strongly Agree). For the present discussion, I will focus only on the metaphysical and ascription statements. The ascription item was similar to response (3) from the experiment by Cohen and Nichols (2010). The metaphysical item, in turn, was designed to elicit intuitions about the ultimate nature of the taste property. Here is one of the responses for the metaphysical item:[14] "In reality, there is an absolute fact of the matter about the taste of the object regardless of how it appears to Alex and Harry and regardless of what they think, say, or do." (ROBERTS; SCHMIDTKE, 2016, p. 2801).

[13] The participants were laypeople with no training in philosophy (n = 330).

[14] I say "one of the responses" because the experimental design included two scenarios: one involving a large disagreement (e.g., moaning vs. beeping) and the other a small disagreement (e.g., honking vs. beeping). The idea was to determine whether participants' responses were influenced by the extent of the disagreement rather than by the properties themselves.

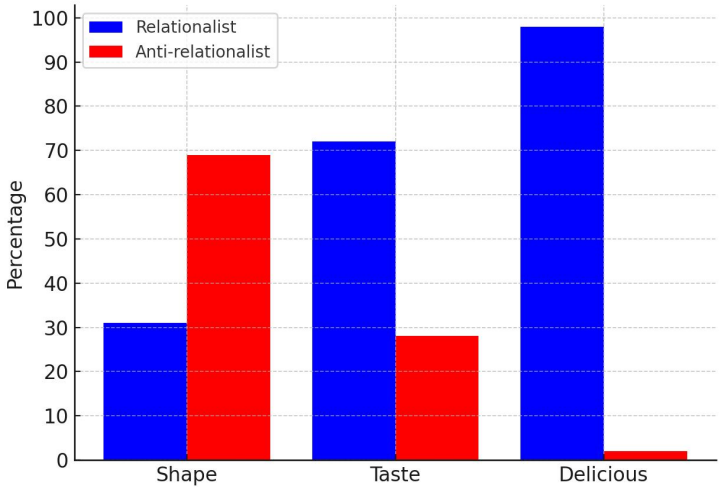

Figure 1 - Mean proportion of relationist versus anti-relationist responses.Graph modified from Cohen & Nichols (2010, p. 224).

The results for both the metaphysical and ascription statements suggest that people's intuitions regarding taste properties are relational. These results were similar to those for the property of likability, which is considered the gold standard for relational properties, just as shape is considered the gold standard for anti-relational properties.

As the experimental evidence from the x-phi of perception makes clear, folk intuitions about taste properties tend to support a relationist view, contrary to what Farkas's analogy with a person suffering from a cold suggests. According to this empirical data, lay judgments typically treat taste as constitutively dependent on a relation (or a complex set of relations) between an object and a perceiving subject. On this view, when a person with a cold remark that "all food seems to have no taste", the utterance is best interpreted as a pragmatic shorthand for something like: "due to my current physiological state, I am unable to instantiate the causal-sensory relation necessary for taste perception". That is, the failure lies not in the object lacking an intrinsic taste property, but in the subject's temporary inability to sustain the perceptual relation constitutive of the taste experience. If Farkas is indeed attempting to support an anti-relationist theory of taste (or by analogy, pain) by appealing to commonsense, then she faces a significant challenge: the empirical record does not corroborate the content of her appeal. However, as I emphasized earlier, my aim is not to refute anti-relationism*per se*, nor to discard entirely the philosophical utility of pre-theoretical judgments. Rather, I intend to critique a specific methodological move: appealing to what I have termed hypothetical commonsense, that is, attributing supposedly shared intuitions to a general

population based solely on private reflection or anecdotal observation. The fact that such appeals are widespread in the literature does not mitigate their epistemic fragility. Indeed, the present evidence demonstrates a divergence between actual empirical judgments and the judgments attributed by Farkas. And even if, by coincidence, Farkas's attribution of folk judgment had matched the empirical findings, this would have been epistemically lucky, not methodologically justified. Mere luck cannot confer justificatory status to a metaphysical claim. For her anti-relationalist thesis about pain or taste properties to be properly grounded, it would need to be supported by epistemically robust and causally appropriate methods, such as those employed in x-phi, cognitive psychology, or neuroscience.

Before concluding, I would like to briefly consider one last objection. One could argue that the arguments I have presented here were *too much from the armchair*, and that the best, or perhaps even the correct, way to support my arguments would be to conduct experiments myself on the intuitions elicited in the cases of the philosophy of consciousness. In other words, the arguments presented here are of the same nature as those I accused of being based on hypothetical common sense. Now, if that is the case, then the very same problems turn back on my arguments. According to this objection, the only legitimate way to support my claims would be to conduct dedicated experiments probing lay intuitions about consciousness. I partially agree with this diagnosis: empirical studies specifically designed to test the claims advanced here would certainly provide further robustness, and I leave such investigations open as a direction for future research. However, I contend that the objection, as typically framed, equivocates on the notion of "armchair argument". To sharpen the issue, let us introduce a distinction.

(I) Cartesian Armchair: An argument *p* is from a Cartesian armchair iff*p* is constructed solely on the basis of *a priori* reasoning and first-person intuitions, without appealing to empirical data (historical example: Descartes' *cogito*).

(II) Naturalist Armchair: An argument *q* is from a naturalist armchair iff*q* is constructed without conducting original experiments but by systematically integrating and interpreting available empirical findings to support a theoretical claim.

On this account, Cartesian armchair arguments rely exclusively on the deliverances of reason, introspection, and/or private intuitions. By contrast, what I shall call naturalist armchair arguments are not experimental in the narrow sense of producing new data, but their justificatory structure is nonetheless constrained by, and responsive to, the body of empirical findings available. In this mode of reasoning, the epistemic warrant of an argument depends on the systematic integration, assessment, and interpretation of extant empirical results, such that its probative force derives not from intuition alone but from its

embedding within an evidentially mediated and scientifically informed framework. The conflation of these two modes of reasoning under the single label "armchair" risks obscuring a crucial methodological distinction. My arguments in this paper fall squarely into the naturalist category because they are not based on private intuition but rather are constructed through the critical analysis of extant experimental data and methodological practices from experimental philosophy and social psychology. That said, even the designation "armchair" may be misleading in this context, as it carries substantial philosophical baggage. Historically, "armchair philosophy" has been employed pejoratively to connote speculative or intuition-driven theorizing detached from empirical reality. To apply the same term to naturalistically oriented work risks a category mistake, since such arguments operate within an evidentially grounded, data-sensitive framework. Thus, the real issue is not whether an argument is "armchair" in some generic sense, but whether it proceeds in the Cartesian mode of *a priori* reasoning or in the naturalist mode of empirically mediated reasoning, and, ultimately, what epistemic resources one brings to bear from it. Accordingly, the objection misfires: the arguments presented in this paper are not "too much from the armchair" at all, but exemplify empirically mediated reasoning in a constructive sense.

6 Conclusion

This paper has argued for an empirically disciplined deployment of the method of cases and against appeals to hypothetical common sense in the philosophy of consciousness. By examining Farkas's defense of exteroceptive pain, I have shown how speculative thought experiments and appeals to commonsense intuitions in remote scenarios can yield epistemically fragile metaphysical conclusions. While such strategies remain widespread in consciousness studies, they face serious challenges from both methodological reflection and empirical data, and are better replaced by a moderate restrictionist stance informed by pragmatic and naturalistic considerations.

I have proposed a moderate restrictionist approach that targets the use of remote scenarios in high-variance philosophical domains. This restrictionism, I argued, is domain-specific and methodologically conservative. Furthermore, I have argued that appeals to common sense require empirical validation, and that philosophers should distinguish between their own introspective projections and the actual judgments of ordinary speakers. Rolla's pragmatic view further supports this conclusion by showing that the reliability of conceptual competences is bounded by the historical and practical contexts in which they were refined, and that when such competences are extended to radically unfamiliar scenarios, their evidential force is systematically undermined.

These criticisms are not intended to undermine the role of *a priori* reasoning in philosophy altogether. Rather, they aim to call attention to the epistemic limits of such reasoning when applied to domains, such as consciousness, in which intuitions have appeared to be culturally variable and theory-laden.

Acknowledgments

I am grateful to Roberto Horácio, Marco Aurélio, Sérgio Farias, and Rodrigo Gouvea for their careful reading and valuable comments on earlier versions of this paper. I also wish to thank two anonymous referees for their insightful suggestions, which greatly improved the final version.

Referências

BENGSON, J. How Philosophers Use Intuition and 'Intuition'. *Philosophical Studies*, v. 171, n. 3, p. 555-576, 2014.

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

BUCKWALTER, W.; STICH, S. Gender and Philosophical Intuition. In: KNOBE, J.; NICHOLS, S. (eds.). *Experimental Philosophy:* Volume 2. New York: Oxford University Press, 2014, p. 307-346.

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

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

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

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

CHUDNOFF, E. *Intuition*. New York: Oxford University Press, 2013.

COHEN, J.; NICHOLS, S. Colours, Colour Relationism and the Deliverances of Introspection. *Analysis*, v. 70, n. 2, p. 218-228, 2010.

COLOMBETTI, G. *The Feeling Body: Affective Science Meets the Enactive Mind*. Cambridge, MA: MIT Press, 2014.

CUTTER, B.; TYE, M. Tracking Representationalism and the Painfulness of Pain. *Philosophical Issues*, v. 21, n. 1, p. 90-109, 2011.

DAVIDSON, D. Knowing One's Own Mind. *Proceedings and Addresses of the American Philosophical Association*, v. 60, n. 3, p. 441-458, 1987.

FARKAS, K. Constructing a World for the Senses. In: KRIEGEL, U. (ed.). *Phenomenal Intentionality*. New York: Oxford University Press, 2013, p. 99-115.

FISCHER, E.; SYTSMA, J. Zombie Intuitions. *Cognition*, v. 215, 2021.

GIRALDO, A. What is the Experimentalist Challenge to the Method of Cases? *Synthese,* v. 205, 2025.

JACKSON, F. Epiphenomenal Qualia. *The Philosophical Quarterly*, v. 32, n. 127, p. 127-136, 1982.

KLEIN, C. *What the Body Commands: The Imperative Theory of Pain*. Cambridge, MA: MIT Press, 2015.

KNOBE, J. Philosophical Intuitions Are Surprisingly Robust Across Demographic Differences. *Epistemology & Philosophy of Science*, v. 56, n. 2, p. 29-36, 2019.

LOCKE, J. *An Essay concerning Human Understanding*. NIDDITCH, P. H. (ed.). Oxford: Oxford University Press, 1975 (1690).

MACHERY, E. *Philosophy Within Its Proper Bounds*. Oxford: Oxford University Press, 2017.

MACHERY, E.; MALLON, R.; NICHOLS, S.; STICH, S. P. Semantics, Cross-Cultural Style. *Cognition,* v. 2, n. 3, p. 1-12, 2004.

MILLIKAN, R. G. *Language, Thought, and Other Biological Categories*. Cambridge, MA: MIT Press, 1984.

MOORE, G. E. *Principia Ethica*. Cambridge: Cambridge University Press, 1903.

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

_____. Swampman, Teleosemantics and Kind Essences. *Synthese*, v. 200, 2022.

PAUTZ, A. The Real Trouble with Phenomenal Externalism: New Empirical Evidence for a Brain-Based Theory of Consciousness. In: BROWN, R. (ed.). *Consciousness Inside and Out: Phenomenology, Neuroscience, and the Nature of Experience*. Dordrecht: Springer, 2014, p. 237-298.

_____. *Perception*. New York: Routledge, 2021.

ROBERTS, P.; SCHMIDTKE, K. A. Relationalism About Perceptible Properties and the Principle of Charity. *Synthese*, v. 193, p. 2779-2803, 2016.

ROLLA, G. Contra Intuições. *Filosofia Unisinos*, v. 22, n. 1, p. 21-28, 2021.

SYTSMA, J. Attributions of Consciousness. *WIREs Cognitive Science*, v. 5, n. 6, p. 635-648, 2014.

TYE, M. *Consciousness, Color, and Content*. Cambridge, MA: MIT Press, 2000.

WATKINS, M.; SHECH, E. *The Metaphysics of Color*. Cambridge: Cambridge University Press, 2025.

WEINBERG, J. M.; NICHOLS, S.; STICH, S. *Normativity and Epistemic Intuitions. Philosophical Topics*, v. 29, n. 1-2, p. 429-460, 2001.

WILLIAMSON, T. *The Philosophy of Philosophy*. 2nd. ed. Malden MA: Blackwell Publishing, 2021.

WITTGENSTEIN, L. *Philosophical Investigations*. Trans. ANSCOMBE, G. E. M.; HACKER, P. M. S.; SCHULTE, J. 4th ed. Malden, MA: Blackwell Publishing, 2009 (1953).

Recebido em: 01/08/2025

Aceito em: 17/09/2025