

## THE DERIVABILITY GAP – A PROBLEM OF INTUITIONS

SONIA MICU<sup>1</sup>

**Abstract:** This paper addresses our difficulty in grasping how something like consciousness could emerge from the processes of our brain, aiming to explain the intuitions that underlie this struggle. I start with an investigation of materialism, revealing that its supporters often have deep-seated dualist intuitions implicit in some of the language they use. I then question whether we can warrant the claim that conscious experience is fundamentally different from non-conscious phenomena with respect to causal powers and causal roles. I identify the derivability gap as the reason behind our intuitive struggles, explaining how these intuitions make it difficult to accept a materialist view of consciousness. Lastly, I explain why we hold onto these intuitions. Instead of seeking a socio-cultural origin for our dualist intuitions about the mind-body problem, I suggest we examine our intuitive grasp of the physical world, arguing that our tendency to see things in binary on/off forms extends beyond our grasp of consciousness. I propose that we consider whether the same simplifying principle that aids our understanding of unobservables might also influence how we conceive consciousness. I conclude that our lack of direct exposure to its complexity may underlie our binary understanding of life and non-life, which we extend to the contrast between material and non-material.

**Keywords:** Consciousness; Hard Problem; Explanatory Gap; Derivability Gap; Illusionism; Meta-Problem

---

<sup>1</sup> Sonia Micu recently defended her doctoral dissertation in Philosophy at the University of Bucharest.

## Introduction

Several theories of consciousness have been developed during the last few years, and despite a clear consensus on which theory has the strongest arguments, materialism seems to be the prevailing view among most contemporary philosophers. In this paper, I will look into the difficulties of intuitively accepting such a view and propose a possible explanation.

David Chalmers (1995) asserted that if any problem can be considered the problem of consciousness, it is the problem of experience. He acknowledged that, when we engage in cognitive activities and perceive the world around us, there is a subjective dimension, a first-personal perspective. Chalmers referred to this subjective aspect as experience. Examples of such experiences include visual sensations like the perception of colors and depth, auditory experiences such as the sound of a clarinet, olfactory experiences like the smell of mothballs, bodily sensations ranging from pain to pleasure, mental imagery, emotional experiences, and the continuous stream of conscious thoughts. All these states share the common characteristic of there being something that it is like to be in them, they are states of experience.

While there is a consensus that experience has a physical basis, there is currently no satisfactory explanation for why and how it emerges from physical processes. This problem of consciousness comes with an 'explanatory gap', a term introduced by Joseph Levine in 1983, according to which an explanation of the physical processes doesn't contribute to our understanding of a subjective experience, such as the way pain feels. As David Papineau (2019) observes, our knowledge of the relation between the physical states and what we subjectively experience doesn't seem to help us overcome the 'dualist' intuition that they are simply different states which accompany each other. Our struggles with understanding consciousness are reflected in what Chalmers calls

" 'phenomenal reports': the things we say about consciousness (that is, about phenomenal consciousness). More specifically, many people make

problem reports expressing our sense that consciousness poses a hard problem. I say things like ‘There is a hard problem of consciousness’, ‘It is hard to see how consciousness could be physical’, ‘Explaining behaviour does not explain consciousness’, and so on.” (Chalmers, 2018, p. 7)

I start by investigating a few comments on materialism which reveal the hard-wired dualist intuitions of both materialists *and* of those arguing against it.

## 2. Materialists as Disguised Dualists

A particular flaw in the materialist stance that has been highlighted by several philosophers, including Saul Kripke, Joseph Levine and David Papineau. Despite arriving at this conclusion using different arguments, they bring light to the same difficulty in overcoming our own intuitive dualism.

As Kripke (1980) construes it, materialism asserts that a comprehensive (and true) understanding of the world can be achieved through a physical description alone. Mental facts are believed to be inherently and necessarily dependent on the underlying physical facts. According to Kripke, no identity theorist has presented a compelling argument against the intuitive view that this strict ontological dependence of mental facts on physical facts may not be accurate. The thorny dialectic surrounding how essences relate to ontological dependence is chronicled in the exchange between Kit Fine and Jessica Wilson in Fine (2020).

Levine (1983) presented a response to Kripke's argument, acknowledging that his own version of the argument does not directly claim the falsity of materialism, making it a less forceful critique compared to Kripke's. Nevertheless, Levine maintained that if his interpretation was accurate, it still presented a challenge to materialism and better captured the discomfort that many philosophers experience in relation to that belief. From Levine's point of view, there seems to be only one practical route for a materialist to confront this dilemma: it involves rejecting the

fundamental intuition upon which the argument is built. Levine suggested that this would require taking a more radical stance in eliminating the concept of qualia, a step that might be too bold for many materialists.

Papineau's (2011) brought in his own proof that even the most materialist of philosophers are, in fact, intuitively resistant to mind-brain identities, and the proof lies in the terminology they use. In his words, "brain processes are standardly said to 'generate', or 'yield', or 'cause', or 'give rise to' conscious states. These expressions are common currency among many thinkers who will insist that they are no dualists. But the phraseology gives the lie to their denial." (Papineau, 2011, p. 12) He uses the example of water-H<sub>2</sub>O identities, in which one doesn't 'generate' or 'give rise' to the other, to show that such words give away a different type of perceived relation between two entities, one closer to the one between fire and smoke. In other words, once we state that brain processes 'give rise' to conscious states, we see them as ontologically different.

Each of these three authors puts forward a criticism of materialism, whether it is in invoking type-identity theory without fully applying its rules, in disregarding qualia or in utilizing a dualist-like language. The dialectic is clearly summarized by Mircea Dumitru (Dumitru 2019, pp. 100–116), along with the lucid diagnostic that much of the literature seems to pertain to shifting the burden of proof from dualists to materialists or the other way around.

Before proceeding to discuss the derivability gap, let us first clarify what relationship a standard version of materialism bears to such a concept. Recently, Papineau (2019), in a response to Chalmers (2018), elaborated on illusionism as part of the conversation on materialism. According to him, since most philosophers are materialists who hold that conscious states are either identical to or fundamentally realized by material states, they are inclined to dismiss the problematic intuitions as false. On their views, people are simply wrong in believing that consciousness possesses non-physical attributes. Undoubtedly, there may be an explanation for why people hold these mistaken beliefs, but those

beliefs are ultimately false. Papineau concludes from this that consciousness is indeed acknowledged to exist by materialist philosophers, but people tend to have numerous misconceptions about its nature. Since Chalmers (2018) used the term "illusionist" to classify anyone who rejected the problematic intuitions, he put standard materialists in the same category as those who claim that consciousness itself is an illusion. For Papineau, the views of the two groups are different. The standard materialist perspective is that the intuitions may be illusory, but that consciousness itself is real. Illusionists not only deny the validity of the intuitions but also deny the existence of consciousness altogether. Papineau sees a dichotomy in Chalmers' classification: either one accepts the problem intuitions and holds a non-physicalist realist view of consciousness, or one rejects the intuitions and is labeled an "illusionist" alongside those who deny the existence of consciousness.

Papineau also claims that the essence of the hard problem, as understood in the context described by Chalmers, revolves around the intuitive claims that consciousness is non-physical. For Papineau, the weak illusionist position provides a viable resolution by asserting that these claims stem from false intuitions, and that the main challenge lies in empirically explaining why these intuitions are so strong. One possible interpretation of Chalmers' stance that Papineau offers is that he might be considering the "hard problem" not as a matter of anti-physicalist intuitions per se, but rather as synonymous with the derivability gap. Papineau notes that weak illusionists do acknowledge the existence of the derivability gap, but they do not perceive it as a problem in its own right.

We can conclude from this that the derivability gap is a relevant concept only for those materialists who embrace a standard materialist view, namely that consciousness is real, of purely physical nature. In this case, addressing the derivability gap can serve as proof, as a solution to the hard problem of consciousness. If one classifies as an illusionist and thus claims that consciousness isn't real, then to them there is no derivability gap and no hard problem of consciousness to solve.

## 2. The Challenge of Identity Statements

Kripke (1980) wrote that

"when Descartes, and others following him, argued that a person or mind is distinct from his body, since the mind could exist without the body. He might equally well have argued the same conclusion from the premise that the body could have existed without the mind. [He added:] let 'Descartes' be a name, or rigid designator, of a certain person, and let 'B' be a rigid designator of his body. Then if Descartes were indeed identical to B, the supposed identity, being an identity between two rigid designators, would be necessary, and Descartes could not exist without B and B could not exist without Descartes." (Kripke, 1980, p. 145)

"The final kind of identity, the one which I said would get the closest attention, is the type-type sort of identity exemplified by the identification of pain with the stimulation of C-fibers. These identifications are supposed to be analogous with such scientific type-type identifications as the identity of heat with molecular motion, of water with hydrogen hydroxide, and the like." (Kripke, 1980, p. 148)

Levine (1983) argues that Kripke's Cartesian argument against materialism rests on two claims:

"first, that all identity statements using rigid designators on both sides of the identitysign are, if true at all, true in all possible worlds where the terms refer; second, that psycho-physical identity statements are conceivably false, and therefore, by the first claim, actually false. [Levine's objective being] to transform Kripke's argument from a metaphysical one into an epistemological one. My general point is this. Kripke relies upon a particular intuition regarding conscious experience support his second claim. I find this intuition important, not least because of its stubborn resistance to philosophical dissolution. But I don't believe this intuition supports the meta-physical thesis Kripke defends – namely, that psycho-physical identity statements must be false. Rather, I think it supports a closely related epistemological thesis." (Levine, 1983, p. 354)

He argues that the reason we can imagine psycho-physical identity to be true is because we lack the epistemological ground which could deny such an intuition. We have an epistemological challenge, or, as Levine calls it, "an explanatory gap" which makes it difficult to know which statements are true. Levine starts his argument with the following assumption:

"To begin with, let us assume that we are dealing with a physicalist type-identity theory. That is, our materialist is committed to statements like:

(1) Pain is the firing of C-fibers.

On Kripke's general theory, if (1) is true at all it is necessarily true. The same of course, is the case with the following statement:

(2) Heat is the motion of molecules.

That is, if (2) is true at all it is necessarily true. So far so good."

(Levine 1983, p.354)

He then proceeds to describe what he calls 'a felt contingency' about these statements. One can indeed imagine a world in which they are false, but this would have to be a logically impossible world, if such identities are deemed to be necessarily true. Therefore, we would need to explain away our felt contingency, even if it appears coherent to us. This seems achievable for (2), since we seem to be able to imagine heat without the underlying motion of molecules but perhaps produced by a different mechanism. Contingency could rather apply to a statement (2') such as:

"The phenomenon we experience through the sensations of warmth and cold, which is responsible for the expansion and contraction of mercury in thermometers, which causes some gases to rise and others to sink, etc., is the motion of molecules" (Levine, 1983, p. 355).

Such a solution would satisfy our felt contingency, but as Levine observes, it would not work for (1). The difference that he remarks between (1) and (2) is that what counts as pain is the experience, the sensation of pain, which makes it impossible to separate the phenomenon

from the sensation, as was the case for heat. Therefore, since in the case of (1) our felt contingency cannot be explained away, the only remaining option is to renounce the truth of (1).

What we can conclude from these arguments is that although materialism claims a type of identity, it isn't a typical one. Comparing identity statements regarding pain and other phenomena doesn't prove helpful. If such an approach had been useful, we could gain an understanding of consciousness by comparing it to other phenomena. The challenge we face is precisely that consciousness isn't like any other phenomenon. Pain and other subjective experiences are a special kind because 'what they feel like' is the phenomenon, not just a by-product.

It bears mention that appreciating the felt intuitiveness of how pain might differ from its physical correlates is typically done while presupposing that holism about phenomenological vocabulary is not well-suited to account for how we use words like "pain". For otherwise it would be questionable to draw inferences from identity statements involving the word "pain" to real identities involving real pain. For discussion, cf. (Quine and Ullian, 2007).

### **3. The Explanatory Power of Functionality**

Let us go back to the claims put forward by Kripke. The difference between claims (1) and (2) is underlined by another more significant difference between the two, one that Levine (1983) puts forward as follows. Nothing of fundamental value is left out from the explanation of the identity of statement (2). The same does not hold for statement (1). Levine successfully captures the explanatory force of statement (2) by formulating the statement (2') above. With it, he shows which mechanism brings about the causal functions of heat, explained by our knowledge of chemistry and physics. For Levine, the two statements (2) and (2') exhaust all there is to be understood about the notion of heat: both its essential nature and its causal role.



A functionalist story would claim that statement (1) does the same for pain. It explains the causal role of pain by referencing an avoidance mechanism that goes into effect when C-fibers are excited by certain nerve endings which are in turn excited by an interaction with the environment such as the penetration of skin with a sharp object. The challenge that Levine sees here is that this explanation does not exhaust the notion of pain. He remarks that the qualitative character of pain is an essential part of the concept of pain, while its connection with C-fiber firing remains mysterious.

Chalmers (1995) also addressed the topic of functional explanation when he pointed out that the distinction between the easy problems and the hard problem of consciousness lies in the nature of the questions they pose.

"The easy problems are easy precisely because they concern the explanation of cognitive abilities and functions. To explain a cognitive function, we need only specify a mechanism that can perform the function. The methods of cognitive science are well-suited for this sort of explanation, and so are well-suited to the easy problems of consciousness. By contrast, the hard problem is hard precisely because it is not a problem about the performance of functions. The problem persists even when the performance of all the relevant functions is explained." (Chalmers, 1995, p. 202)

I argue that the conversation on functionality raises two issues. First of all, the assumption that 'what it feels like' is a defining part of the pain, without which we cannot fully explain it, also starts from a dualist perspective, that there is a physical and a separate non-physical nature to pain. This is an intuition still to be validated. *As with other ideas from this paper, one can either start from the intuition and attempt to explain it, or start from explaining the phenomenon and show that an intuition is not justified. If we decide to use intuitions, we must name them for what they are.*

Second, if we say that other processes can be explained through their functional role while for consciousness such an explanation isn't possible, we might seem to assume that qualia have no causal role in

phenomenal consciousness. I propose we consider a scenario in which what an experience is like plays a functional role in that experience, a role perhaps still to be uncovered. Such an assumption might lead us to the conclusion that an agent might not be fully functional without qualia, that philosophical zombies might lack some functionalities by lacking consciousness.

#### **4. An Alternative View on Dualist Intuitions**

Even if given a full scientific account of the functionality of any phenomenon attributed to consciousness, we would still lack an explanation of the experience associated with that phenomenon. This is known as the explanatory gap. Chalmers (1995) acknowledges Levine's (1983) "explanatory gap" between cognitive functions and conscious experience, a gap that needs bridging in order to gain a comprehensive understanding of consciousness. He agrees that merely describing the functions does not provide an explanation for the subjective experience that accompanies them. While conscious experience may play a cognitive role, any functional explanation alone falls short in accounting for the phenomenon. It is possible that exploring functions in depth might lead to insights that contribute to understanding experience, but such discoveries would be additional explanatory rewards. Chalmers suggests that the conventional methods of cognitive science and neuroscience, developed to explain cognitive functions, are insufficient for addressing the hard problem of consciousness.

We might argue that when it comes to the topic of consciousness, the dualist intuition is everpresent. Papineau (2019) argues that Chalmers (2018) already presupposes, by his phrasing, that consciousness is of a non-physical nature, when he describes the hard problem as "why and how do physical processes in the brain give rise to conscious experience?" (Chalmers, 2018, p. 6) and, later, when he introduces a category of gap intuitions as those by which "there is an explanatory gap between physical processes and consciousness" (*ibid.*, p. 12). Papineau explains

that, if one entity 'gives rise to' another, they must possess distinct ontological statuses. For instance, fire gives rise to smoke, but H<sub>2</sub>O does not give rise to water. To Papineau, these psychological responses are brute intuitions that stem from a preexisting commitment to dualism as the alternative explanation. People's initial conviction that the mind is separate from the brain leads them to be naturally perplexed by the capacity of brain processes to generate conscious phenomena. They are dissatisfied with physiological explanations and wonder why the brain 'gives rise to consciousness'.

Papineau proposes that we consider what he calls the "derivability gap" as the underlying cause for the hard problem. Along with its associated explanatory gap, it is, in his view, the evident explanation for the perplexity we experience in regards to consciousness. As Papineau remarks, Chalmers has consistently advocated that the 'hard problem' and the 'explanatory gap' are both caused by the absence of a priori derivability. He has been arguing that the 'easy problems' of consciousness are 'easy' precisely due to their reliance on functional concepts specifying roles. Processes such as learning or memory can be accounted for because we have an understanding of the functional roles that they play, which enables us to identify corresponding physical mechanisms. Chalmers has also emphasized that the 'hard problem' emerges precisely because phenomenal states are not subject to a priori analysis, and the apparent "explanatory gap" stems from our incapacity to a priori derive the existence of phenomenal facts from our knowledge of physical mechanisms.

Papineau (2020) agrees with Chalmers that explaining 'problem intuitions' is key to a satisfactory account of consciousness. He comments:

"According to the mainstream view, we think of salt as the stuff that is white, crystalline, granular, with a distinctive taste, that dissolves in water, and is found in the oceans. Now imagine someone who has a fully detailed account of the physical make-up of the world, in terms of the distribution of matter, arrangement of elementary particles, the deployment of fields, and so on. In principle, such a person could arguably put this knowledge together with their prior conceptual grasp

of salt to figure out that salt must be NaCl, on the grounds that NaCl is the stuff that fits the conceptual requirements for salt – white, crystalline, ... However, we can't do this with pain, say, or with visual experiences of red" (Papineau, 2020, p. 18)

We think of pain in terms of the feelings it generates, not in terms of some role it plays. And so, connecting physical facts with the phenomenon of pain is something that doesn't come easy to us. We can't derive mind-brain identities and this is what creates a feeling of puzzlement about them. We conceive scientific properties and conscious properties differently.

I conclude from Papineau's approach that, if the explanatory gap is an epistemological gap, the derivability gap doesn't seem to be of epistemic kind, but of a deeper intuitive nature. I argue that such a gap that can never be closed. A scientific explanation of how and why consciousness arises from physical processes *might solve the explanatory gap for us. It might, however, have no impact on the derivability gap*, if this gap is as deeply routed in the way we conceive the world as Papineau states.

Is our intuitive dualism a proof against materialism as one might assume from Kripke's argument or rather a chance to solve the hard problem of consciousness as in Chalmers' hypothesis? For Levine (1983), the fact that this deep-seated intuition about our subjective experiences proves to be so resistant to philosophical dissolution shows that the enduring puzzle of the mind/body problem will stay with us for as long as its corresponding intuition does. Papineau invokes an inability to free ourselves from "an implicit commitment to dualism" (Papineau, 2011, p.8) as the cause for our feeling that something is left unexplained in this mind-brain problem. We have a strong belief that our pain cannot be just some C-fibers firing, that our conscious states cannot be reduced to brain states. For Papineau, there would be no explanatory gap if we only overcame our intuitive resistance and accepted mind-brain identity. Instead, the dualist nature of our thoughts on mind and brain makes it difficult to identify some phenomenal kind with a material kind. In his view, the presence of the dualist intuition does not pose an argument

against materialism itself. He suggests that the difficulty faced by materialism is *primarily psychological rather than theoretical*.

Papineau's main argument is that the intuitive implausibility of materialism does not pose a problem for the philosophical position itself. Materialism, according to Papineau, is a well-supported and coherent standpoint. The fact that many individuals find materialism difficult to believe is not sufficient to discredit its validity, since many truths are challenging to accept. He suggests that if the intuitive implausibility of materialism presents a problem, it is a problem for materialists to address, rather than a fundamental flaw in materialism. Materialists should recognize and examine the influence of dualistic intuitions on their thinking, adjusting their perspectives accordingly. He acknowledges that some may view the dualist intuition as evidence against materialism and expect materialists to explain or dismiss this intuition by demonstrating why it persists despite being false. However, *he refuses to concede that the dualist intuition inherently supports dualism or undermines materialism*. Instead, he asserts that the superiority of dualism's explanation for the dualist intuition over materialism's explanation should not be assumed without thorough investigation and evaluation. While acknowledging the significance of the dualist intuition and the desire to explain it, he argues against taking it as conclusive evidence against materialism or in favor of dualism and emphasizes the need to critically assess whether dualism can provide a more compelling explanation for the psychological phenomenon of the dualist intuition than materialism can.

I propose we look into an explanation of this intuitive dualism. We may notice the pattern we have of seeing the world through a binary approach when we think of how we understand living organisms. We are incapable of intuitively accepting a correspondence between some unnoticeable chemical processes and what is happening at macro level in a living organism. The only way in which we realized that life isn't something of a different nature, separated from chemistry, was to create very simple chemical structures and observe the passing from inorganic to organic: the same chemical components arranged in a certain structure that we call 'life', molecules organized in a way that forms 'living'

organisms. We learned that a complex animal is nothing but a reapplication of the same principle, that what we call 'life' is a form of organization of molecules.

### A Final Word

The derivability gap is based on our incapacity to intuitively accept that something like our consciousness can derive from the processes in our brain. We might blame this on our intuitive dualism. If our intuitive dualism is standing in the way of our understanding of consciousness, it is instrumental to know the origin of this intuition. I offer a possible route. Rather than look into a socio-cultural source for our mind-body dualist intuition, I propose we look into how we perceive the world around us. Our binary view seems to go beyond the problem of consciousness.

Everything we learned from science about the continuum between inert matter and living organism, which teaches us of the gradual chemical changes that make the transition to life in an organism, doesn't seem to liberate us from the binary view of life and non-life. We cannot observe with a naked eye these micro-phenomena happening in the world, and so we use a binary classification. I propose that we look into whether the same simplifying principle that has helped us make sense of the world we couldn't micro-observe and understand is a principle hidden in our perception of what consciousness is. Not being directly exposed to this complexity might create our binary conception of life/non-life that we have taken further and associated with the difference between material and non-material. If we cannot accept that life isn't of a different separate nature then these processes, but rather 'derives' from chemistry, it is understandable that we also cannot *intuitively* accept that consciousness 'derives' from brain processes.

If this assumption is explored and validated by further research, I believe it can offer an answer to the problem of phenomenological reports and why we find consciousness so puzzling, if not dissolve the hard problem itself.

## REFERENCES

- Chalmers, D. (1995). Facing up to the Problem of Consciousness, *Journal of Consciousness Studies*, Volume 2, pp. 200-219.
- Chalmers, D. (2018). The Meta-Problem of Consciousness, *Journal of Consciousness Studies*, Volume 25 (9–10), pp. 6-61.
- Dumitru, M. (2019). *Lumi ale gândirii. Zece eseuri logico-metafizice*. Iași: Polirom.
- Fine, K. (2020). Comments on Jessica Wilson's "Essence and Dependence". In: *Metaphysics, Meaning and Modality: Themes from Kit Fine* (pp. 471–475). Edited by Mircea Dumitru. Oxford, US: Oxford University Press. <https://doi.org/10.1093/oso/9780199652624.003.0033>
- Kripke, S. A. (1980). *Naming and Necessity*, Harvard University Press.
- Levine, J. (1983). Materialism and Qualia: The Explanatory Gap, *Pacific Philosophical Quarterly*, Volume 64, pp. 354-361.
- Papineau, D. (2011). What Exactly Is the Explanatory Gap?, *Philosophia*, Volume 39, pp. 5–19.
- Papineau, D. (2019). Response to Chalmers' 'The Meta-Problem of Consciousness', *Journal of Consciousness Studies*, Volume 26, Numbers 9–10, 2019, pp. 173–181(9).
- Papineau, D. (2020). The Problem of Consciousness. In *The Oxford Handbook of the Philosophy of Consciousness*. Oxford, UK: pp. 14–36.
- Quine, W. V., and J. S. Ullian. (2007). *Țesătura Opiniilor*. (Romanian edition of *Web of Belief*). Translated by Mircea Dumitru. Pitești, Romania: Paralela 45.