



# Intuition-Driven Navigation of the Hard Problem of Consciousness

Krzysztof Sękowski<sup>1</sup> · Wiktor Rorot<sup>1</sup>

Accepted: 17 February 2021 / Published online: 27 February 2021  
© The Author(s) 2021

## Abstract

The discussion of the nature of consciousness seems to have stalled, with the “hard problem of consciousness” in its center, well-defined camps of realists and eliminativists at two opposing poles, and little to none room for agreement between. Recent attempts to move this debate forward by shifting them to a meta-level have heavily relied on the notion of “intuition”, understood in a rather liberal way. Against this backdrop, the goal of this paper is twofold. First, we want to highlight how the ontological and epistemological status of intuitions restricts the arguments in the debate on consciousness that rely on them. Second, we want to demonstrate how the deadlock in those debates could be resolved through a study of a particular, “positive” kind of intuitions. We call this approach “The Canberrish Plan for Consciousness” as it adopts elements of the methodological “Canberra Plan”.

## 1 Introduction

While the “hard problem of consciousness” (Chalmers 1996) - the question of why and how the physical processes going on in the brain become conscious - is central to the current study of consciousness, the discussion surrounding it has become somewhat stale. There are two well-defined, opposed camps of realists and eliminativists of all sorts, and almost no room for agreement between them. Recently, as a way out of this clinch, David Chalmers has proposed the “meta-problem of consciousness”, the problem of why we think that there is a hard problem at all (Chalmers 2018: 6). What the meta-problem does is it introduces a new level into the debate - one on which the two camps can agree on the formulation of the issues in question, and work together to develop a viable research strategy. This in result could provide support to the different standpoints at the “ground floor” - the standard level - of the debate.

- 
- ✉ Krzysztof Sękowski  
k.sekowski@uw.edu.pl
  - ✉ Wiktor Rorot  
w.rorot@uw.edu.pl

<sup>1</sup> Institute of Philosophy, University of Warsaw, Warsaw, Poland

Chalmers' formulation of the meta-problem makes extensive (and liberal) use of the notion of "intuition", underscoring the importance of intuitive reasoning and arguments in the discussion of the hard problem. However, his reliance on the notion of intuition is extremely problematic, as we will argue in what follows. "Intuition" is not a transparent notion, and contemporary debates on the question of its ontological and epistemological status are far from settled. In fact, as we will show in this paper, a particular stance with regard to the status of intuition fixes one's views on consciousness. In result, such a formulation of the meta-problem is begging the question.

There is, however, a different way of tackling this issue. By focusing on positive intuitions, that is investigating particular (philosophical, as well as everyday) views on conscious experience, we may implement a sort of a "Canberrish Plan for Consciousness" "(for Canberra Plan see e.g. Braddon-Mitchell and Nola 2009 and section 5 below). In this way, the problem of the ontological and epistemological status of intuitions can be avoided.

In what follows, we will summarize the discussions of the hard problem (section 2) and the Chalmers' proposal of the meta-problem programme (section 3). Next, we will argue that Chalmers' programme fails to navigate the dispute between illusionists and realists. In particular, we will show that the nature of intuitions determines the possible solutions to both the meta-problem and the hard problem (section 4). Finally, we will argue that there is hope that comes from studying positive intuitions about phenomenal experience according to the proposed "Canberrish Plan for Consciousness", introduced in section 5.

## 2 Realism Vs Illusionism

Philosophy of consciousness - and consciousness studies in general - has been largely structured around the famous "hard problem" of consciousness, formulated by David Chalmers (1996). The hard problem can be stated as follows: "Why is all the information processing in the brain accompanied by an experienced inner life?" Why do some physical states become conscious, appear in a direct way to the subject, with an associated "something it is like" for the subject to be in that state? Chalmers distinguished this from the "easy" problems which are susceptible to a reductionist, functional explanation, and as such can be directly studied using empirical methods. The hard problem, on the other hand, asks about the phenomenal, first-person, subjective qualities of the experience of the world, which are left out by functionalist accounts. This incredibly influential proposal has shaped most of the current debate on consciousness, even though some prominent scholars, such as Daniel Dennett, believe that the "so-called problem is a chimera" and a distraction from the real hard question: "and then what happens?": "once some item or content 'enters consciousness', what does this cause or enable or modify?" (Dennett 2018: 1).

The major division in contemporary consciousness studies lies in the researchers' attitude towards the hard problem (see e.g. Frankish 2016). Realists take the explanatory gap highlighted by the hard problem as impassable via empirical research and attempt to otherwise account for phenomenal experience. They typically claim that phenomenal consciousness is a real "something" going on in the brain (or rather – in the mind) beyond what our best functional theory can explain. Realism can take many

forms, from classical, Cartesian-inspired substance dualism, via more sophisticated forms of dualistic views (advanced e.g. by Chalmers 1996), to less popular views such as Russelian neutral monism and various panpsychist approaches, which view mental states as a basic “ingredient” of reality (see e.g., Goff 2017).

Realism, for a long time, has been opposed by eliminativism, a view that held that the idea of phenomenal states, similar to many other concepts from folk psychology, is simply wrong and the hard problem is misconstrued altogether: there is nothing in the mind above and beyond of what we can explain with reference to functional categories. The explanatory gap simply does not exist.

In recent years, however, a more nuanced approach has been slowly gaining popularity. Illusionism has been on the table since at least the 1980s (Dennett 1991; the name has been proposed much later, by Frankish 2016), however for a long time it remained a minority view. It differs from earlier eliminativist approaches in that it takes the (illusion of) consciousness seriously. In this view, the “special”, “ineffable”, “subjective” character of phenomenal experience, is real inasmuch as the Necker cube “really” changes its position as we’re looking at it. To stay with this simile: the instability of the cube is a valid explanatory target, however, once we explain that there are no intrinsic depth cues in the picture and the brain switches between two compatible interpretations of the shape nothing “above and beyond” remains to be accounted for. The change is *explained away*. Similarly, illusionists argue that even though the (illusion of) phenomenal experience is real, once we explain all “easy” problems of consciousness, there will be nothing else to add to the explanation of this illusion. In fact, under this view, the explanatory gap itself is an illusion.

This debate, while organized around the hard problem, for a large part has been focused on the question of whether or not the hard problem is in fact meaningful. In result, there is little room for possible agreement between the proponents of realism and illusionism, and the discussion of phenomenal consciousness slowly becomes circular. Without assigning any blame, it seems uncontroversial to state that there is currently an impasse between the two views.<sup>1</sup>

### 3 Meta-Problem as a Way Forward

A recent paper by Chalmers (2018), offers a new way of addressing the question posed by the hard problem. Chalmers aims to move the discussion of the nature of consciousness to a meta-level. He argues that “The Meta-Problem of Consciousness” may help bridge the gap between the two camps and provide a novel research programme addressing the issues raised by the hard problem of consciousness.

Chalmers formulates the meta-problem as follows: “The meta-problem is (...) the problem of explaining why we think consciousness is hard to explain” (Chalmers 2018: 6). In short, “the meta-problem of consciousness is (...) the problem of explaining (...) problem reports.” (Chalmers 2018: 7). He points out that at least some people (including

<sup>1</sup> We believe that it is important to disclose at this stage that the authors of this paper themselves differ in their views on consciousness: one (KS) is more realistically oriented, while the other (WR) tends towards illusionism. Our later proposal for navigating the dispute between realists and illusionists, laid out in section 5, seems satisfactory from both points of view.

himself) have dispositions to make certain judgments that are in fact descriptions of different kinds of difficulties one encounters whenever they try to grasp the nature of consciousness, similar to the claim “consciousness is hard to explain”. The dispositions to make this kind of judgments are dubbed by Chalmers “problem intuitions”, and the verbalizations of these intuitions are called “problem reports”. Chalmers offers several examples: “There is a hard problem of consciousness”, “It is hard to see how consciousness could be physical”, “Explaining behavior does not explain consciousness” (Chalmers 2018: 7), which are obviously spelled out in a language of a professional philosopher. However, we can imagine – as Chalmers suggests by pointing out to the body of work on “intuitive dualism” – non-philosophers formulating sentences like “There is something mysterious about me having conscious experience of colour”, “It is impossible to scientifically study consciousness” or “Okay, so this is what the brain does, but how do I experience X”?<sup>2</sup> Problem intuitions are interesting as they belong to functional consciousness, they are a part of behavior, and as such are “easy” to explain. At the same time they seem to be telling us something about the phenomenal consciousness and how it is experienced. Chalmers argues that they constitute a special subclass of the easy problems, as explaining where the problem reports come from may throw some light on the hard problem - independently of whether one is a reductionist or not. Hence, explaining how the problem intuitions arise could serve as a proxy for an explanation of the hard problem.

In short, there are two reasons why the meta-problem could provide a new research paradigm, convincing to both realists and illusionists. First, the meta-problem is formulated in a language that can be accepted by both realists and illusionists. Problem intuitions and reports replace controversial “qualia” or “phenomenal experience” as explanatory targets. This strategy is compatible with both realists’ reliance on and illusionists’ reluctance towards qualia and phenomenal properties.

Second, the meta-problem provides novel constraints on possible solutions to the hard problem, allowing to arbitrate among available options with respect to how much they agree with registered problem intuitions.

However, in the next section we will show that although the idea of moving the dispute on the nature of consciousness to a meta-level is promising, the program proposed by Chalmers ultimately fails. The critique boils down to the fact that the ontological and epistemological status of “problem intuitions” depends on the status of intuitions in general in a way that ties respective answers to the question about intuitions to respective solutions to the hard (and meta) problem of consciousness.

## 4 Problems with (Problem) Intuitions

The aim of this section is to show that the meta-problem implicitly begs the question. First, we will show how the relation between problem intuitions and phenomenal consciousness could be understood. Then, in subsections 4.1–4.3, we will argue that solving the hard problem of consciousness by explaining problem intuitions is already begging the question, since the way we understand the notion of *intuition* would lead directly to either some kind of realistic or illusionistic account. We will introduce three

<sup>2</sup> We don’t have to refer to our imagination - reports mentioned here are actual examples of reports produced by non-philosophers prompted by one of the authors (in private conversations).

ways of understanding the nature of intuitions that lead to particular positions regarding the nature of consciousness. Two of them lead to some kind of illusionism (section 4.1, 4.2), and one to a realistic account (4.3).

Chalmers doesn't reflect on the status of intuitions, stating only very generally that problem intuitions are dispositions to make problem judgments and problem reports (at the same time being distinct from e.g. phenomenal beliefs - Chalmers 2018: 46). However, a good understanding of what are intuitions in general, and problem intuitions specifically, turns out to be crucial for the purpose of the meta-problem approach.

Meta-problem programme relies on the claim that intuitions are somehow related to phenomenal consciousness. However, justifying that claim proves quite complicated. In fact, there are three possible relations between the problem intuitions and phenomenal consciousness (Chalmers 2018: 48): *meta-problem nihilism* or *correlationism*, which posits no causal relationship between intuitions and phenomenal states; *meta-problem realizationism*, which claims that phenomenal consciousness is the primary cause of problem intuitions, and hence depends on a realistic view of consciousness; and *weak* and *strong illusionism*, according to which the problem intuitions are either illusory (*strong illusionism*) or are aimed at certain other, lower- or higher-level mental processes (*weak illusionism*). Both in strong and weak illusionism the primary cause of problem intuitions is shared with the processes that give rise to the illusion of phenomenal consciousness.

Note that, at the beginning of investigating the problem intuitions, we have to, at least roughly, bear in mind their ontological, and epistemological/methodological<sup>3</sup> status. It means that we start with some assumptions about what intuitions are, what is their role in cognition and in philosophical reflection. These issues are connected with each other in such a way that some ontological stances entail epistemological ones and vice versa.

In result, once we accept any particular attitude towards intuitions, the solution to the meta-problem, as well as the hard problem, follows directly from this stance, rather than from any other philosophically interesting considerations about consciousness. In our view, solutions to the meta-problem of consciousness are in fact views about the content and reliability of problem intuitions.

In the remainder of this section we will substantiate these claims, showing more directly the logical interconnectedness between intuitions and consciousness.

#### 4.1 Strong Illusionism

Some philosophers argue that intuitions are simply beliefs or dispositions to believe. It seems that this indeed is Chalmers' standpoint, when he states that an intuition is a "disposition to judge and report" (Chalmers 2018: 12; this view is also shared by other philosophers engaging in the meta-problem discussion, e.g. Clarke-Doane 2019; Schwarz 2019). We can call proponents of this view "reductionists" or "eliminativists" about intuitions. This nomenclature calls attention to the fact that these philosophers usually deny that intuitions have any epistemically privileged role, especially in

<sup>3</sup> We assume that epistemological and methodological status of intuitions are strictly related to one another. If someone agrees for example that intuitions are reliable a priori insights it would entail that we could rely on intuitions in philosophical methodology, and treat them as evidence. If someone thinks that intuitions are just as reliable as any other beliefs it would entail that intuitions should not have any privileged methodological status.

philosophical methodology. They do not set any specific conditions that dispositions to believe or mere beliefs must meet to count as intuitions. This kind of approach is represented by e.g., Timothy Williamson (2007), Peter van Inwagen (1997) or David Lewis (1983). Its deflationist character is best illustrated in what Lewis writes about intuition:

"Our 'intuitions' are simply opinions; our philosophical theories are the same. Some are commonsensical, some are sophisticated; some are particular, some general; some are more firmly held, some less. But they are all opinions..."  
(Lewis 1983: X)

To determine whether this stance about the nature of intuitions implicitly begs the question on the meta-problem, we should consider possible epistemological stances coherent with the scrutinized ontological account. This view is often accompanied by the methodological claim that intuitions are in fact totally redundant and we should avoid them in philosophical practice. But this methodological claim can be weakened, and still coherent with the core "reductionist" view. If intuitions are mere beliefs - as reliable and as questionable as any other beliefs - we can still find some space for them in our philosophical inquiry. The term "intuition" may be redundant, but we cannot refuse that ordinary beliefs do play some role in philosophical methodology.

In this way, beliefs could be treated as a common ground or a starting point in theorizing (see Williamson 2007: 242). Therefore, by studying intuitions about consciousness, and problem intuitions in particular, we could find a common ground for investigating the nature of the folk concept of consciousness.

These are two epistemological views coherent with the stance that intuitions are mere beliefs or mere dispositions to believe. The first one is that we should abandon intuitions in philosophical enterprise all together. Having done so, the meta-problem cannot be adopted, since it rests on studying a particular kind of intuition - *problem intuitions*. Such claim is held e.g. by Rosenthal (2019). However, as mentioned, according to the second available epistemological view, we could treat intuitions as a common ground in philosophical disputes. In the case of the meta-problem, we could treat problem intuitions as a common ground for discussing the hard-problem of consciousness. In that case, however, we should ask about how well justified are problem intuitions, as they are mere beliefs or mere dispositions to believe.

Note, that problem intuitions in such a case are as reliable as any other beliefs, e.g. as reliable as the belief that the Earth is flat (which is false), that Pluto is a dog (which is true), or that all combustible bodies contain an element called "phlogiston", which has negative mass and is released during combustion (which is false). This view is in fact supported by the increasing amount of research indicating that beliefs about consciousness are in fact culture-dependent (Irvine 2019; Sytsma and Ozdemir 2019; Yetter-Chappell 2019) and in case of scientists working on phenomenal experience - theory-laden (Lau and Michel 2019). In this case, it is hard to see how we can assign any special role to intuitions about consciousness, specifically - to problem intuitions on which the Chalmers' meta-problem rests.

One could argue that intuitions about phenomenal states differ somewhat from other kinds of intuitions in that they report our beliefs about our own mental states. Introspection seems to be the most plausible mechanism of their origin.

According to traditional accounts of introspection, we have a privileged access to our own mental states (Schwitzgebel 2010; Schwitzgebel and Cushman 2012). This privilege is understood in various ways. Some philosophers (Burge 1988; Papineau 2002; Chalmers 2003) maintain that self-ascriptions are always true in a self-fulfilling way. Others (Brentano 2015 [1874]; Chisholm 1969) argue that properties of our own mental states are self-presenting, so that knowledge is non-inferential.

However it is disputed that self-reports are indeed more reliable than other typical beliefs. Arguments against the claim that we have privileged access to our own mental states are mostly based on empirical evidence. We can point to the whole range of research that has been accumulated since 1970s showing that people have poor knowledge about their own mental states or processes that underlie behaviour (Nisbett and Wilson 1977; Nisbett and Bellows 1977; Cherniak et al. 1983; Wegner and Wheatley 1999; Mele 2001; Wilson 2002; Johansson et al. 2005).

This is not a place to resolve whether the proponents or opponents of the privileged access are right (and to what extent). Nevertheless, the following holds: if we accept that intuitions are ordinary beliefs, then we have to reject privileged access. Otherwise the beliefs about our own mental states are special because of privileged access and do not qualify as *mere beliefs* anymore. Rather, they would constitute a special kind of beliefs (we will discuss what would happen with this kind of special beliefs in section 4.3.).

Therefore, we cannot claim both that intuitions are mere beliefs, and that *problem intuitions* are better justified than any other beliefs. What does it mean for the metaproblem programme? There are two possible answers. The first one is that problem intuitions could be true, however there is nothing special in their epistemic status. The second option is that they are just misguided beliefs.

In the first case there is no reason to adopt the meta-problem programme in order to explain or solve the hard problem of consciousness. It is because in Chalmers' enterprise problem intuitions serve as a proxy for the explanation of the hard problem of consciousness because of their special status. However if they are just as reliable as any other beliefs, which could be as possibly true as false, then there is no gain in grounding the possible explanation of the hard problem of consciousness on the explanation of the problem intuitions.

It should be obvious what happens if we accept the second view, according to which intuitions are just misguided beliefs. It would mean that there is nothing special or surprising about some people being convinced that their conscious experience is non-physical or hard to explain (just as there is nothing "special" or "uncanny" about some people being convinced that the earth is flat). This is precisely what the illusionists claim regarding the hard problem of consciousness. There is nothing special with what we call problem reports, just as there is nothing special about our belief that there is something called "phenomenal consciousness." The question of how we arrive at this particular belief may be of some interest. However, in line with the claims of strong illusionists, the view that intuitions are ordinary beliefs and that there is no privileged access to our mental states, dissolves the meta-problem (and in turn - the hard problem as well). There is simply nothing

else to add to the explanation of why problem intuitions arise in some people in the first place.

In sum, once we agree that intuitions are mere beliefs, we need to abandon the whole meta-problem enterprise all along or else we turn out to be illusionists. There does not seem to be a third way out of this dilemma.

## 4.2 Weak Illusionism

Another ontological view on intuitions is to hold that intuitions are special beliefs or special dispositions to believe. For some proponents of this account, the special aspect of relevant beliefs or dispositions to believe can be described as follows: someone has an intuition that *p* solely on the basis of competence with the concepts involved in *p* (Ludwig 2007: 135), or someone has an intuition that *p* merely on the basis of understanding *p* (Sosa 1998).

This perspective can be accompanied by a methodological claim that intuitions serve either as a starting point or as a touchstone for philosophical theories. Some philosophers who agree with this view, regard conceptual analysis grounded in intuitions as the basic philosophical method. It is partly because they expect that intuitions should reveal some necessary truths *a priori* (e.g., BonJour 1998; Ludwig 2007). This could include some necessary truths about consciousness and in result, these special beliefs could form a much stronger foundation for the investigation of phenomenal experience. Their special status makes them less defective and less dependent on philosophically irrelevant “background noise”, when compared to beliefs formed through experience.

However, this view limits intuitions to *conceptual intuitions*. They make use of a specific concept of consciousness. First of all, it could be that the concept is a widespread, “general public” one. This is the option that seems preferable to any investigation of consciousness on the basis of intuitions. This view is also entertained by Chalmers in his meta-problem argument, as he discards that the “problem intuitions” arise solely for philosophers and claims that they are in fact widespread among the lay people (Chalmers 2018: 15). However, if it is such a folk concept, then the historical, cultural and socioeconomic factors come into picture (as highlighted by Irvine 2019; Sytsma and Ozdemir 2019; and Yetter-Chappell 2019) and the possibility of making strong, conceptual arguments of the kind imagined by some vanishes, as the intuitions turn out to be fallible, in the manner already discussed in the previous section.

On the other hand, we could consider some kind of an improved, philosophical concept of consciousness, which relies on the folk one, but slightly changes its meaning in the course of conceptual analysis. In such case, however, it is highly unlikely that everyone has the required kind of conceptual competence, because otherwise this philosophical concept would not be different from the folk one. Furthermore, in this case the special status cannot be regarded as coming into picture based on the privileged access to our mental states. And since not everyone has the kind of conceptual competence required, then how can we verify who does? Philosophers? Consciousness researchers? While some philosophers are likely to endorse this answer (e.g., Ludwig 2007: 149), the perspective of understanding the nature of consciousness based on intuitions alone loses its appeal in this context (see e.g. Lau and Michel 2019).

What's more, accepting that intuitions are a special subclass of beliefs or dispositions to believe entails a form of illusionism - higher order illusionism.

Before we move to the argument note that in this approach, beliefs are obviously different from conscious experiences, as for example we may hold beliefs that we are unaware of or are not actively attending, and in general they do not require any phenomenal component.

The argument explaining why accepting a stance that intuitions are special (conceptual) beliefs would entail weak illusionism, or at least exclude realism, can be presented as follows:

Suppose that it is possible to be a realizationist (and thereby realist) and hold the view that intuitions are special beliefs that are justified by competence or understanding of the subject matter. Now, consider the meta-problem realizationism that Chalmers presents as the main way for a realist to tackle the meta-problem (Chalmers 2018: 42). Realizationism claims that consciousness directly gives rise to (or is a primary cause of) intuitions about phenomenal experience. If so, and if intuitions are special beliefs resulting from conceptual competence, then consciousness would have to play a direct causal role in the conceptual beliefs about consciousness itself. In other words, phenomenal experience should play a direct causal role in forming conceptual beliefs about consciousness. But then either it is the case that:

- There is no additional process required and we have competence with regard to consciousness simply in virtue of having conscious experience.
- Or there is an additional process which makes the difference between “normal” beliefs and intuitions about consciousness (this could be the process of introspection, if we would agree for privileged access, but then see section 4.3).

As for the first option – it is highly implausible to assume that it is true. Note that it would entail that our *conceptual* beliefs about consciousness follow *directly*, so without any intermediate process, from our phenomenal experience. But several processes take part in acquiring conceptual beliefs in a linguistic form about any given subject – consider at least the processes responsible for language processing.

The view which holds that conceptual beliefs about consciousness follow directly from our phenomenal experience could come from the undetected difference between *phenomenal character* and *phenomenal concept* (see Tye 2003: 91). Phenomenal character is a quality of our phenomenal experience. According to some philosophers – especially the proponents of qualia, we could have direct access to phenomenal character. However, phenomenal concepts are the result of processing the phenomenal character of our experience. Therefore, phenomenal concepts come indirectly from phenomenal experience, because several processes are needed to establish these concepts.

The idea expressed above is in fact well seated. In Kantian jargon we can point out that to form concepts our sensations alone are insufficient. We have to use both the faculty of understanding and receptivity to formulate any judgment of perception. This idea is expressed by Kant in his famous quote “Thoughts without content are empty, intuitions without concepts are blind.” (Kant 1787/1998: B75).

Therefore accepting the stance that there is no additional process required for appearance of problem intuitions and we have competence with regard to

consciousness simply in virtue of having conscious experience can not be matched with the view according to which intuitions are conceptual beliefs.

Now, let us turn to the option that there is an additional process which makes the difference between “normal” beliefs and intuitions about consciousness. If this is the case, then the primary causal role is played by this additional process. But this is contrary to the assumption of realizationism. In fact this is in line with the claims of higher-order illusionism since the additional process is responsible for our problem intuitions, and not phenomenal experiences themselves. Hence, if we try to build a theory of consciousness based on intuitions understood as special conceptual beliefs, only the claims of higher-order illusionism remain viable. “Only creatures with certain introspective models will be phenomenally conscious” (Chalmers 2018: 43) and therefore only those creatures will have intuitions about consciousness.

### 4.3 Realism

The final ontological possibility is to hold that intuitions are *sui generis* mental states that cannot be reduced to other mental states such as beliefs or dispositions to believe. This *sui generis* state can be approximately characterized as a state in which some proposition seems true (Bealer 1998; Pust 2000), or less precisely, as a state that comes with a peculiar phenomenology that attends the experience of seeing that some proposition is true (Plantinga 1993: 105–6). Under this view, intuitions and beliefs are independent. One can have an intuition that *p* without believing that *p*, like in the Müller-Lyer illusion, in which we can have an intuitive sense that one of the arrows is longer than the other without believing this to be the case.

Intuitions thus understood are methodologically often taken as a starting point of philosophical reflection which reveals necessary truths, providing a priori justification (Bealer 1998; Pust 2000: 39). Such a perspective offers the strongest conceptual foundations for the work on consciousness that departs from intuitions, including the meta-problem.

Unfortunately, accepting that intuitions are *sui generis* mental states and referring to intuitions about consciousness in an attempt to understand the nature of phenomenal experience also begs the question. If we regard intuitions as states of some proposition seeming to be true, or a mental state associated with some peculiar phenomenology, by definition we assume that there exists a phenomenal property that differentiates intuitions from beliefs. Thereby we deny the possibility of reducing this *sui generis* phenomenal state to any other mental state or process.

In sum, the acceptance of an ontological view according to which intuitions are *sui generis* states differing from other mental states by their phenomenology entails that there exists an irreducible phenomenology within the mind. Hence, employing intuitions to explain the hard problem of consciousness can lead only to the view according to which phenomenological experiences are irreducible.

Alternatively, a similar problem arises if we accept the claims discussed previously, namely that intuitions are mere or special beliefs, and at the same time accept privileged access to one’s own mental states. Under this view, we can get to know something about mental states directly by observing our “mental life”, without the need to infer mental properties from behaviour. However, if this direct observation is to constitute knowledge - and hence if it can be reliably referred to in the discussion of the meta-

problem - the justification associated with these beliefs has to be accounted for. If the knowledge in case of privileged access is direct, it would seem that this justification can only come from an accompanying phenomenal state of seeing that some proposition is true, similar to the one that makes intuitions *sui generis* mental states.

In a nutshell, the uniqueness of intuitions as a distinct class of mental events hinges on the phenomenology they come with: that of directly seeing the truthfulness of some proposition. Once we accept this view, there is no room for eliminativism or illusionism, as we would have to point out to specific mental processes this phenomenology reduces to, disassembling the metaphysical uniqueness of intuitions in the process.

It seems that the approach to intuitions which is the most promising for a consciousness researcher does not in fact allow for a meaningful study of the problem, as the answer is assumed already at the very beginning. Furthermore, specifically in relation to the meta-problem introduced above, this view of intuitions will lead to a rejection of the Chalmers' claim that the meta-problem is a special, "gateway" easy problem, as it turns out to be only a different face of the hard problem.

#### 4.4 Meta-problem's Failure

As we have previously mentioned, the introduction of the meta-problem of consciousness was supposed to serve as a new field of discussion between the proponents of opposing views of consciousness, on which they could agree at least with how the problem is defined. However, by the extensive use of the notion of intuition, and specifically the central concept of "problem intuitions", the meta-problem shifts the discussion to a dispute over intuitions themselves. However, this does not seem to be much of an improvement for the ongoing debates between illusionists and realists, as one's views on the nature of intuition fix the response to the questions about the nature of phenomenal experience, as we have argued above.

### 5 Studying Positive Intuitions and the Canberra Plan

Despite the issues with the meta-problem programme discussed in the previous section, Chalmers' main idea to move the debates on consciousness to the meta-level, and place the study of intuitions in the center of such a meta research programme is still promising. In this section we will show how we believe it is possible to engage with Chalmers' ideas in a way that is not subsumed under the critique we've laid out in the previous section. We will call this proposed approach "The Canberrish Plan for Consciousness".

The main idea is that the meta-research should focus on the very concept of "consciousness" rather than specifically on the meta-problem of consciousness. This means that the researcher should grapple with all kinds of intuitions about consciousness, instead of studying only what Chalmers dubbed "problem" intuitions. In other words, the focus should be on the study of positive intuitions (i.e. intuitions about what consciousness *is*) instead of negative ones (i.e. intuitions about what consciousness *is not* or why consciousness *cannot be explained* by physicalist science). The methodology of this kind of approach can be transplanted from another intuition-centered methodological project: The Canberra Plan (see e.g. Braddon-Mitchell and Nola

2009). Such an approach would be able not only to clarify already existing debates on consciousness, but also to advance it beyond the entrenched oppositions of realists and illusionists.

The Canberra Plan is an influential methodological approach proposed and developed mainly by David Lewis and Frank Jackson (Lewis 1970, 1972; Jackson 1998).<sup>4</sup> According to this programme, metaphysical investigations should start from studying intuitions and platitudes about the analyzed concept. This is the first stage of the conceptual analysis. The result is a list of characteristics of the concept we are interested in. The second stage, then, is to find out empirically what scientific entities satisfy these characteristics (all of them or at least a majority included in such a list).

In the case of the Canberrish Plan for Consciousness, the proposed research programme also consists of two stages. First, we should collect intuitions and platitudes about consciousness, focusing on lay people, and carefully paying attention to a representation of diverse backgrounds, cultures and worldviews. It is possible, given that Chalmers is right that problem intuitions are widespread, that in the course of conducting the research there would be some intuitions of the form “consciousness is hard to explain” collected. But most likely the majority would be of a positive form, that is ascribing certain characteristics to consciousness rather than denying them, including statements such as “conscious experience is subjective and perspectival”, or perhaps such as “consciousness is inconstant and stressful”.<sup>5</sup> In result, we could compile a list of characteristics that are ascribed to consciousness.

In the second stage we envisage a slight departure of the Canberrish Plan from its forebear. The Canberra Plan is plainly a naturalistic account (Kornblith 2016: 155). The proper program for navigating the dispute on the nature of consciousness should however allow also for non-naturalistic theories of consciousness. The Canberrish Plan for Consciousness does not presuppose a naturalistic framework of the analysis’ results. We propose that in the second stage the search for theories capable of accounting for the characteristics of consciousness should be extended beyond the accounts vindicated by contemporary science. The second stage would therefore consist of either finding empirically what kind of entities satisfy the list of characteristics from the first stage or determining a priori what kind of entities would satisfy that list. This seems to avoid the risk of rejecting non-physical (or more broadly - anti-reductionist) theories of consciousness at the outset.

Limiting ourselves to the study of problem - or negative - intuitions, as Chalmers does in his formulation of the meta-problem, we remain within the confines of current debates on consciousness. Most importantly - within the confines of the opposition of realism and illusionism, which is damning for the program of studying intuitions, as we have shown previously. This limitation stems from the fact that negative intuitions are primarily not about consciousness itself, but about the hard problem: “why consciousness cannot be explained in such and such a way” or “why consciousness cannot be characterised so and so”. Through the theory-ladenness of the questions, this “such and such” and “so and so” are presuppositions that limit possible explanations to the ones

<sup>4</sup> It is worth noting that Chalmers himself has advocated the Canberra Plan with regard to metaphysics in general (see Chalmers 2012).

<sup>5</sup> As some Buddhist traditions claim, see Samyutta Nikaya in (The Connected Discourses of the Buddha 2000: SN 22.79).

entertained by either realists or illusionists. However, with the inclusion of positive intuitions, the study of intuitions escapes the boundaries of this opposition and avoids begging the question on the status of intuitions themselves. This approach allows for the possibility that none of the extant accounts is correct or even on the right track, hence the current impasse. The novel accounts may be compatible with different views of intuitions and in this way, the Canberrish Plan for Consciousness avoids the criticisms laid out previously.

We see multiple possible outcomes of this kind of study. Maybe some of the negative intuitions held only by philosophers would disappear in the process. Maybe this research would indicate which properties of the concept of consciousness affect the reception of scientific claims and explanations and make them unconvincing. It could also throw light on the pre-theoretical entanglements of problem intuitions specifically, highlighted by some commentators of Chalmers' meta-problem proposal. For example, Rosenthal (2019) argues that problem intuitions might be only intuitions of people with specific theoretical (or pre-theoretical) background. This view hinges on the question whether the concept of consciousness underlying problem intuitions is the same concept that lies at the heart of the scientific discourse on consciousness. What's more, Rosenthal points out that intuitions can be elicited by the structure of wording of the question posed (Rosenthal 2019: 196), and argues strongly that they should not be regarded as constraints on explanation of consciousness. Similar point is also made by Wierzbicka who argues that problem intuitions will depend on the metalanguage in which they are studied (Wierzbicka 2019: 263).

It is also possible that the Canberrish Plan for Consciousness will provide some insight into why people believe that there is in fact a hard problem of consciousness. It could turn out that for some consciousness is not a concept from the same category as properties or entities which can be explained by physical science. Consciousness could be e.g. regarded as a property of a person, where "person" is understood in terms of some religious or metaphysical assumptions that are incompatible with physical vocabulary. In this case, this research would highlight a difference between the very concepts used by some lay people and most of the researchers immersed in current discussions about consciousness. Translating the results about one of those concepts onto another would in fact constitute a category mistake.

Note that the Canberrish Plan for Consciousness is very inclusive in terms of the understanding of the notion of "intuition". It goes with any ontological approach to intuitions, as well as with any position towards the methodological and epistemic status of intuitions. The only exceptions are accounts according to which intuitions should be entirely abandoned in any intellectual enterprise due to the lack of clarity about what they are (such views could be assigned to Cappelen (2012) or Williamson (2007)).

In the way that the proposed programme centers the views of consciousness held by lay people and strives to be compatible with both eliminativist and anti-reductionist accounts, it is reminiscent of Daniel Dennett's heterophenomenology (e.g. Dennett, 1991, 2003).<sup>6</sup> Similarly to the proposed Canberrish Plan for Consciousness, the goal of heterophenomenology was to "compose a catalogue of *what the subject believes to be true about his or her conscious experience*" (Dennett 2003: 20, italics original). There is however a major difference. In Dennett's approach this catalogue of beliefs was to be extracted from the reports concerned with some particular experiences. This is

<sup>6</sup> We are thankful to the anonymous reviewer for bringing those similarities to our attention.

highlighted by the metaphor of a generic psychological experiment in which the subject is asked to report their train of thoughts. In this regard heterophenomenology happens *exactly on the level* of phenomenal consciousness. In turn, in the Canberrish Plan we are interested in the concept of consciousness the interrogated subject has: platitudes and intuitions about what consciousness in principle is, rather than what the subject's current experience consists of. That is precisely why the postulated programme (similarly to Chalmers' meta-problem) shifts the discussion to the *meta level*. Furthermore, we do not wish to make as strong statements as Dennett does, in claiming that "The total set of details of heterophenomenology, plus all the data we can gather about concurrent events in the brains of subjects and in the surrounding environment, comprise the total data set for a theory of human consciousness. It leaves out no objective phenomena and no subjective phenomena of consciousness." (Dennett 2003: 20). The Canberrish Plan for Consciousness at its first stage explicitly leaves out all phenomena of consciousness in order to construct an account of what falls under the name of "phenomena of consciousness" from the bottom up - directly from folk beliefs. The important difference between the Canberrish Plan and both naturalistic Canberra Plan and Dennet's heterophenomenology is that in Canberrish Plan it is not determined whether an object falling under the concept of "phenomena of consciousness" has to be empirically known (as in Canberra Plan) or correlated with events in the brain (as in Dennet's heterophenomenology). Phenomena of consciousness could meet these features, but they don't have to. In sum, studying positive intuitions in line with the Canberrish Plan for Consciousness in lieu of the meta-problem programme introduced by Chalmers could move the existing debates forward. The account proposed here could shed a light on the actual intuitions about the concept of consciousness, exposing possible differences in understanding this central notion between those who recognize the hard problem of consciousness as a problem, and those who do not. Moreover, this approach opens up the possibility of establishing some new theories of consciousness, and breaking the deadlock between realists and illusionists.

## 6 Conclusion

Much of the discussion on consciousness finds itself in an impasse. Illusionists and realists often do not agree about the very formulation of the "hard problem of consciousness" - the problem around which most of the discussion revolves. The meta-problem of consciousness has been proposed by Chalmers (2018) as a way out of that impasse. Unfortunately, his program fails. The notion of "intuition", specifically the problem intuitions, are central to his research. However, the ongoing debates on the ontological and epistemological status of intuitions must be taken into account. In fact, accepting any one of the three most popular ontological approaches to the nature of intuitions begs the question. If we accept that intuitions are either simply beliefs or some special beliefs, we are led to accept the claims of, respectively, strong or weak illusionism. At the same time, if we accept that intuitions are independent of beliefs and constitute a *sui generis* mental states, only realism remains a viable option. In fact, this shows that any intuition-driven approach to the hard problem of consciousness is doomed.

However, the central idea of Chalmers' meta-problem, namely that of moving the debate on consciousness to a meta-level of sorts, through reference to intuitions can be salvaged. If we abandon the focus on the hard problem, and the negative intuitions that come along with it, we may attempt to implement a "Canberra Plan for Consciousness": a qualitative study of what intuitions about qualities of consciousness are widely shared among people (philosophers and non-philosophers alike). Such a study would be able to provide not only a novel, valid explanatory target for theories of consciousness - or at least with a corroboration of the importance of features which consciousness researchers already bring under scrutiny. It could also hint at novel approaches to explaining the nature of consciousness. In result, it could also wade in on the debates about the hard problem and the explanatory gap it postulates.

**Acknowledgements** We are extremely grateful to the Reviewer from Review of Philosophy and Psychology for their suggestions which highly helped us improve the paper. We would like to also thank Joanna Komorowska-Mach, Mel Andrews, Antonina Jamrozik and the members of the Student Association for Philosophy of Mind at the University of Warsaw for many useful comments and help in editing the manuscript. This work was partly funded by the Ministry of Science and Higher Education (Poland) research Grant DI2017 001347 (PI: Krzysztof Sękowski) as part of "Diametowy Grant" program. Both authors contributed equally.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Bealer, G. 1998. Intuition and the autonomy of philosophy. In *Rethinking intuition: The psychology of intuition and its role in philosophical inquiry*, ed. M.R. DePaul and W.M. Ramsey. Lanham: Rowman & Littlefield.
- BonJour, L. (1998). In defense of pure reason: A rationalist account of a priori justification.
- Braddon-Mitchell, D., and R. Nola, eds. 2009. *Conceptual analysis and philosophical naturalism*. Cambridge: MIT Press.
- Brentano, F. 2015. *Psychology from an empirical standpoint*. Abingdon: Routledge.
- Burge, T. 1988. Individualism and Self-knowledge. *Journal of Philosophy* 85 (11): 649–663. <https://doi.org/10.5840/jphil1988851112>.
- Cappelen, H. 2012. *Philosophy without intuitions*. 1st ed. Oxford: Oxford University Press.
- Chalmers, D. 1996. *The conscious mind: In search of a fundamental theory*. New York: Oxford Univ. Press.
- Chalmers, D. 2003. The nature of narrow content. *Philosophical Issues* 13 (1): 46–66. <https://doi.org/10.1111/1533-6077.00004>.
- Chalmers, D. 2012. *Constructing the world*. Oxford: Oxford University Press.
- Chalmers, D. 2018. The Meta-problem of consciousness. *Journal of Consciousness Studies* 25 (9–10): 6–61.
- Cherniak, C., R. Nisbett, and L. Ross. 1983. Human inference: Strategies and shortcomings of social judgment. *The Philosophical Review* 92 (3): 462. <https://doi.org/10.2307/2184495>.
- Chisholm, R.M. 1969. On the Observability of the self. *Philosophy and Phenomenological Research* 30 (1): 7. <https://doi.org/10.2307/2105917>.
- Clarke-Doane, J. 2019. Undermining belief in consciousness. *Journal of Consciousness Studies* 26 (9–10): 34–47.

- Dennett, D.C. 1991. *Consciousness explained*. Boston: Back Bay Books.
- Dennett, D.C. 2003. Who's on first? Heterophenomenology explained. *Journal of Consciousness Studies* 10 (9–10): 19–30.
- Dennett, D.C. 2018. Facing up to the hard question of consciousness. *Philosophical Trans Royal Soc B: Biol Sci* 373 (1755): 20170342. <https://doi.org/10.1098/rstb.2017.0342>.
- Frankish, K. 2016. Illusionism as a theory of consciousness. *Journal of Consciousness Studies* 23 (11–12): 11–39.
- Goff, P. 2017. *Consciousness and fundamental reality*. Oxford: Oxford University Press.
- Irvine, E. 2019. Explaining variation within the Meta-problem. *Journal of Consciousness Studies* 26 (9–10): 115–123.
- Jackson, F. 1998. *From metaphysics to ethics: a defence of conceptual analysis (reprinted)*. Oxford: Clarendon.
- Johansson, P., L. Hall, S. Sikström, and A. Olsson. 2005. Failure to detect mismatches between intention and outcome in a simple decision task. *Science* 310 (5745): 116–119. <https://doi.org/10.1126/science.111170>.
- Kant, I. 1787/1998. *Critique of pure reason*. P. Guyer & A. W. Wood, Eds. Cambridge, New York: Cambridge University Press.
- Kornblith, H. 2016. Philosophical Naturalism. In H. Cappelen, T. S. Gendler, & J. Hawthorne (Eds.), *The Oxford Handbook of Philosophical Methodology* (Vol. 1). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199668779.013.8>.
- Lau, H., and M. Michel. 2019. A socio-historical take on the Meta-problem of consciousness. *Journal of Consciousness Studies* 26 (9–10): 136–147.
- Lewis, D. 1970. How to define theoretical terms. *The Journal of Philosophy* 67 (13): 427. <https://doi.org/10.2307/2023861>.
- Lewis, D. 1972. Psychophysical and theoretical identifications. *Australasian Journal of Philosophy* 50 (3): 249–258. <https://doi.org/10.1080/00048407212341301>.
- Lewis, D. 1983. *Philosophical Papers Volume I*. New York: Oxford University Press. <https://doi.org/10.1093/0195032047.001.0001>.
- Ludwig, K. 2007. The epistemology of thought experiments: First person versus third person approaches. *Midwest Studies in Philosophy* 31 (1): 128–159. <https://doi.org/10.1111/j.1475-4975.2007.00160.x>.
- Mele, A.R. 2001. *Self-deception unmasked*. Princeton: Princeton University Press.
- Nisbett, R.E., and N. Bellows. 1977. Verbal reports about causal influences on social judgments: Private access versus public theories. *Journal of Personality and Social Psychology* 35 (9): 613–624. <https://doi.org/10.1037/0022-3514.35.9.613>.
- Nisbett, R.E., and T.D. Wilson. 1977. Telling more than we can know: Verbal reports on mental processes. *Psychological Review* 84 (3): 231–259. <https://doi.org/10.1037/0033-295X.84.3.231>.
- Papineau, D. 2002. *Thinking about consciousness*. Oxford : New York: Clarendon Press ; Oxford University Press.
- Plantinga, A. 1993. Warrant and proper function. In *New York*. Oxford: Oxford University Press.
- Pust, J. 2000. *Intuitions as evidence*. Routledge: Place of publication not identified.
- Rosenthal, D. 2019. Chalmers' Meta-Problem. *Journal of Consciousness Studies* 26 (9–10): 194–204.
- Schwarz, W. 2019. From sensor variables to phenomenal facts. *Journal of Consciousness Studies* 26 (9–10): 217–227.
- Schwitzgebel, E. 2010. Acting contrary to our professed beliefs or the Gulf between Occurrent judgment and dispositional belief: Acting contrary to our professed beliefs. *Pacific Philosophical Quarterly* 91 (4): 531–553. <https://doi.org/10.1111/j.1468-0114.2010.01381.x>.
- Schwitzgebel, E., and F. Cushman. 2012. Expertise in moral reasoning? Order effects on moral judgment in professional philosophers and non-philosophers. *Mind & Language* 27 (2): 135–153. <https://doi.org/10.1111/j.1468-0017.2012.01438.x>.
- Sosa, E. 1998. Minimal Intuition. In *Rethinking intuition: The psychology of intuition and its role in philosophical inquiry*, ed. M.R. DePaul and W.M. Ramsey. Lanham: Rowman & Littlefield.
- Sytsma, J., and E. Ozdemir. 2019. No problem: Evidence that the concept of phenomenal consciousness is not widespread. *Journal of Consciousness Studies* 26 (9–10): 241–256.
- The Connected Discourses of the Buddha. 2000. *Tr. Bhikkhu Bodhi*. Somerville: Wisdom Publications.
- Tye, M. 2003. A theory of phenomenal concepts. In *Minds and persons*, ed. A. O'Hear, 91–106. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511550294.006>.
- Van Inwagen, P. 1997. Materialism and the psychological-continuity account of personal identity. *Noûs* 31: 305–319. <https://doi.org/10.1111/0029-4624.31.s11.14>.
- Wegner, D.M., and T. Wheatley. 1999. Apparent mental causation: Sources of the experience of will. *American Psychologist* 54 (7): 480–492. <https://doi.org/10.1037/0003-066X.54.7.480>.
- Wierzbicka, A. 2019. From 'consciousness' to 'I think, I feel, I know': A commentary on David Chalmers. *Journal of Consciousness Studies* 26 (9–10): 257–269.

- Williamson, T. 2007. *The philosophy of philosophy*. Malden: Blackwell.
- Wilson, M. 2002. Six views of embodied cognition. *Psychonomic Bulletin & Review* 9 (4): 625–636. <https://doi.org/10.3758/BF03196322>.
- Yetter-Chappell, H. 2019. Idealization and problem intuitions: Why no possible agent is indisputably ideal. *Journal of Consciousness Studies* 26 (9–10): 270–279.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.