

## Hard, Harder, Hardest

Katalin Balog

Penultimate draft

In this paper, I will discuss three problems concerning phenomenality<sup>1</sup>. The first two problems have been dubbed “The Hard Problem”<sup>2</sup> and “The Harder Problem”<sup>3</sup>. The third problem has received less attention and I will call it “The Hardest Problem”. However, I do not really mean to indicate a competition among these problems as to which is the most difficult. They are all hard (perhaps equally so) and I think that they cannot really be solved separately. Their solution has to come in a piece, and that’s part of what makes all of them so hard.

Together the three problems present, I suggest, a particularly difficult challenge to those philosophers who are both physicalists<sup>4</sup> and phenomenal realists,<sup>5</sup> and agree with dualists that there

---

<sup>1</sup> Throughout the paper, when I talk about phenomenal states, I mean states for which there is, in Thomas Nagel’s expression (1974), “something it is like” to have them. “Phenomenality” is the feature of such states that constitutes what it is like to have them. In the literature “consciousness” is often used interchangeably with “phenomenality; but since the term “consciousness” is used in a number of different senses as well, I will stick to “phenomenality.

<sup>2</sup> See Chalmers 1996, pp. xii-xiii.

<sup>3</sup> See Block 2002.

<sup>4</sup> By physicalism I mean the metaphysical view that the world’s fundamental ontology is physical and the best account of that ontology is provided by fundamental physics. According to contemporary physics, this ontology consists of particles, strings and fields of various types that occupy space-time (or bear spatio-temporal relations to one another) and possess a limited number of quantitative states (mass, charge, electromagnetic potential and so on). Everything else, whatever other entities and states there are, is composed out of and realized by configurations of this fundamental physical ontology. Physics also claims that there are only a few fundamental dynamical and perhaps non-dynamical laws that govern the structure of space-time and evolution of its occupants. All macroscopic or special science laws, causal relations, probabilities are ultimately derived from the laws of fundamental physics and the arrangement of fundamental physical entities.

<sup>5</sup> By “phenomenal realism” I mean that there are phenomenal states. Our concepts of these – phenomenal concepts – possess both first person and third person applications. First person applications have a cognitive immediacy while third person applications are mediated by perception (typically of behavior). I take it for granted that I know in my own case that phenomenality (and various phenomenal states) exists but for other people and creatures there is room for various degrees of skepticism.

is an explanatory gap involving phenomenality.<sup>6</sup> This is not to say that dualists and Russellian monists<sup>7</sup> don't have their own problems – it is just that these problems don't arise in the same form for them. The problems concerning phenomenality that beset dualism and Russellian monism are at least as hard as these but I confine myself here to the problems that arise within a physicalist metaphysics.<sup>8</sup> My aim here is to spell out the relations between these three problems and then to explore how they appear from the perspective of an approach originally championed by Brian Loar (1990/1997) that has been quite fruitful in so far as the first two problems are concerned.<sup>9</sup>

Loar's approach – dubbed “the phenomenal concept strategy” by Stoljar (2005)<sup>10</sup> – attempts to explain the various special and puzzling features of phenomenal states in terms of their relationship to the concepts we apply to them. It is useful here to invoke the distinction between first person and third person concepts; while third person concepts, such as TABLE<sup>11</sup> or ELECTRON do not depend on how the subject is situated or related to the referent, first person concepts do. Some examples of first person concepts are indexicals, such as HERE or NOW or I. Phenomenal concepts are special, subjective, first person concepts whose canonical applications are based on introspection of the subject's own mental states as in the thought I HAVE ALREADY HAD THIS FEELING IN MY KNEE where the notation “THIS FEELING” indicates a phenomenal concept

---

<sup>6</sup> Levine (2001) observes that there is an explanatory gap between phenomenally and physically characterized facts. I explain the gap in Section 1. Those who think there is a gap reject (as I do) analytic functionalist accounts of phenomenal concepts and hold that zombies are conceivable.

<sup>7</sup> A version of Russellian monism that has recently been revived (Chalmers 1996, 2002, and 2017; Lockwood 1998; Stoljar 2001; and Strawson 2003, and 2006) is the view that the most fundamental states are both mental (or “proto-mental”) and physical in that they possess a physical dispositional nature and a mental categorical nature. It is an interesting issue how this view fares with regard to the three problems but that is a topic for another paper.

<sup>8</sup> For a discussion of the problems for dualism see Balog 2005.

<sup>9</sup> Balog 2012a, 2012b, Block 2002.

<sup>10</sup> A similar proposal by Scott Sturgeon (1994) appeals to the special epistemology of phenomenal states.

<sup>11</sup> I will indicate concepts and thoughts with capitalization.

that I formed on the basis of introspecting current sensations in my knee.

Phenomenal concepts are widely thought to be special and unique among concepts in that they refer – in some further specifiable way – directly yet substantially to phenomenal experience. Later, in Section 2, I will have more to say about their exact nature. Of course, not all applications of phenomenal concepts are introspective applications. Clearly, one can token a concept that refers to pain without literally experiencing pain – these can be called “indirect phenomenal concepts” – as when I reply to my dentist’s question by “I am not in pain” or when I see someone stub their toe and think THAT HURTS. Because only phenomenal concepts applied in introspection pertain to the arguments in this paper, for the rest I will ignore these indirect phenomenal concepts.

The basic idea of the phenomenal concept strategy is that dualism seems plausible not because it is true but merely because of structural features of phenomenal concepts. Recently some philosophers who advocate a physicalist approach to both the Hard and the Harder problem (e.g., Papineau 2002, 2003) have called into question whether what I call the Hardest Problem – the problem of explaining, within the bounds of physicalism, how phenomenal concepts can refer determinately to physical or functional states of the brain – can be solved. I propose that the phenomenal concept strategy is capable of handling not just the first two problems but the Hardest Problem as well.

Here is a first pass at stating the three problems.

The *Hard Problem* is a *metaphysical*, and *explanatory problem* concerning the nature of phenomenal states. There is an expectation that if physicalism is true then there ought to be an account of how physical states (most likely, brain states) give rise to phenomenality. Failing that, there ought to be a physicalistically acceptable explanation of why it is *not* possible to give such an explanation.

The *Harder Problem* is *epistemological*. The problem is that if physicalism is true then all facts supervene on physical facts, including facts about phenomenality and so it is natural to expect that, given enough physical information, one can know whether another being has phenomenal states and, if he/she/it does, the character of her phenomenal states. But it seems that one cannot know this. The concern is especially acute with respect to beings that are physically not much like humans. It seems one couldn’t really know if a physically different creature, say, a dog, a fish, a Martian, or very

sophisticated robot, with which humans share certain functional similarities, overlaps with humans phenomenally one way rather than another way or not at all.<sup>12</sup> If physicalism is true then one ought to be able to explain in physical terms why our knowledge of others' phenomenal states seems limited in this way even though they involve perfectly objective physical facts. This problem is closely connected with the Hard Problem and I will propose an account of phenomenal concepts that goes some distance in handling both.

The *Hardest Problem* is a problem about *reference*. Both the Hard and the Harder Problems presuppose that one's subjective phenomenal concepts refer determinately – modulo vagueness – to real, objective states that can be instantiated in minds other than one's own. This assumption is equivalent to the plausible idea that there is a matter of fact – even if it is not possible to know – about whether a phenomenal concept that applies to me also applies to another creature; whether we share our phenomenology. The Hardest Problem is the problem of explaining how this could be so in a physicalist framework. It has proven very hard to provide an account of how a phenomenal concept can refer to a determinate physical or functional state.

The problem is especially difficult given two facts about phenomenal concepts. One is that they are not physically or functionally defined; they have a special direct way of referring. And the other is that it is quite plausible that there will be *several distinct* physical and/or functional states that are correlated with any given phenomenal state. What would determine that a phenomenal concept refers to one of these to the exclusion of others?

The difference between these alternatives is vast. Take a phenomenal state P that is correlated with the neurophysiological state N and functional state F<sub>1</sub> and F<sub>2</sub>.<sup>13</sup> N might be only instantiated by humans, whereas F<sub>1</sub> might be instantiated by fish as well and F<sub>2</sub> might be instantiated by humans, fish *and* robots. On the assumption of determinacy, there are three distinct scenarios corresponding to different metaphysical and semantic facts. If our phenomenal concept C of P refers to N (N then

---

<sup>12</sup> Block 2002 uses the label 'Harder Problem' for a whole cluster of issues concerning the epistemology of phenomenal states that goes beyond the scope of this paper. I am going to use the label only for the issue outlined above; consequently, my usage is slightly different from Block's.

<sup>13</sup> N, of course, realizes both F<sub>1</sub> and F<sub>2</sub>.

being identical to P) that means P only occurs in humans but not in fish or robots. If *C* refers to  $F_1$  then P occurs in both humans and fish but not in robots. And finally, if *C* refers to  $F_2$  then P occurs in humans, fish *and* robots. It seems to be a perfectly factual matter which one of these scenarios obtains. But it is hard to give a physicalist account that corresponds to one or another of these alternatives as it is hard to think of a physical way in which reference could be established to one of N,  $F_1$  and  $F_2$  to the exclusion of the others.

This was a first, rough pass at formulating these three interlocking problems for physicalism. I will now proceed as follows. In Section 1, I go into more detail about them. In Section 2, I take a look at them in the light of the phenomenal concept strategy. This approach is promising for physicalism in that it provides a plausible account of how phenomenal states can be physical even though there is an irremediable explanatory gap (Hard Problem), why it might be impossible to find out whether other creatures have phenomenal states even if phenomenality is purely physical (Harder Problem); and it can also deal with many other philosophically perplexing features of phenomenality, like privileged access, infallibility, etc. Given this promise, I want to see whether, from a physicalist point of view, we are at the point where all of the pieces of the puzzle fall into place. In Section 3, I tackle the Hardest Problem for the physicalist project. I propose that the supposed referential indeterminacy of phenomenal concepts is a misconception and that the phenomenal concept strategy is capable of handling the Hardest Problem as well.

## 1. Three problems

### *The Hard Problem*

The Hard Problem is a metaphysical and explanatory problem for physicalists concerning the nature of phenomenal states. On the physicalist assumption the question arises how physical states of the brain give rise to phenomenality. Block 2002 (p. 423) puts it like this: “Why is the scientific basis of a phenomenal state the scientific basis for that state rather than another or rather than a nonphenomenal state?” Since this question seems to lack a perspicuous answer, the physicalists have a “hard problem” on their hands.

The consequence is various challenges to physicalism. One is the *Zombie*<sup>14</sup> Argument. David Chalmers<sup>15</sup> has argued that if phenomenality were physical, phenomenal propositions (that is, propositions involving phenomenal concepts) would have to be *a priori derivable* from the complete physical truth about the world.<sup>16</sup> However, he argues that phenomenal propositions are not so derivable. This is made evident by the fact that zombies are conceivable. The conclusion is that physicalism is false. In my opinion, this argument is not sound,<sup>17</sup> but it challenges physicalists to explain why zombies are conceivable.

Another argument against physicalism is based on explanatory considerations. As Levine observes<sup>18</sup>, there is an explanatory gap between neuro-physiological (or functional) and phenomenal descriptions. Discovering which neuro-physiological (or functional) states are correlated with a given kind of phenomenal state (say, experiencing red) doesn't explain why or how phenomenal states arise at all. Once there was also an explanatory gap between, for example, ordinary talk of transmission of traits of parents to their offspring, and physical and biochemical descriptions. But that gap has mostly been bridged by genetic theory and molecular genetics. We, or at least molecular geneticists, have a pretty good understanding of how biochemical processes can provide the mechanisms that underlie the transmission of traits from parents to children. Once we understood the molecular biology of genes, we would understand why a particular gene codes for such and such proteins and, given further knowledge about the organism, we would understand why that in an

---

<sup>14</sup> Zombies are (conceivable) creatures that are physically exactly like us but have no phenomenal states at all.

<sup>15</sup> See Chalmers 1996, 2009.

<sup>16</sup> The claim is a little more complicated. It is that every non-indexical truth is derivable *a priori* from the full physical description of the world and the laws and a proposition to the effect that this is the full physical description and laws.

<sup>17</sup> See Balog 1999, especially with regard to Chalmers (1996). His 2009 formulation of the argument is harder to refute, but, as I will show later, can be equally defused using the phenomenal concept strategy.

<sup>18</sup> Levine 2001, pp. 76-80.

organism codes for eye color. The case of phenomenality is different. Those scientific theories that might come forth seem utterly unexplanatory.

Take the simpler example of H<sub>2</sub>O and water. Even if you deny (*pace* advocates of the conceivability arguments) that the statement 'H<sub>2</sub>O=water' is a priori derivable from the full physical description of the world, the hypothesis that, for example, H<sub>2</sub>O is not water is hard to make sense of.<sup>19</sup> Once we have all the micro-physical information about H<sub>2</sub>O and understood and gathered all relevant macroscopic information about water, we can close the explanatory gap between water and H<sub>2</sub>O, no matter whether this is construed as a priori derivation or as posteriori theorizing.

The case of phenomenality is different. Though the search is still ongoing, it is not impossible that one day the neural correlate of phenomenality will be found. Neurophysiological state N or functional state F is the neural correlate of phenomenal state P if there is at least a nomological connection between them: if N or F is coextensive with P in all creatures natural and artificial in (at least) all nomologically possible worlds.<sup>20</sup> The physicalist claims that the relationship between P and N or F is actually metaphysical rather than nomological; that P is identical to either N or F. However, the hypothesis that neurophysiological state N, or functional state F is not identical to phenomenal state P is clearly intelligible, and an open question, for *any* N or F, no matter how much we know about the actual neurophysiology. And so, if physicalism is true, we will never be able to understand why.

---

<sup>19</sup> (Block and Stalnaker (1999) discuss the possibility of 'ghost water' – a non-physical kind that exists side by side with being composed of hydrogen and oxygen atoms and has all the same causal roles as the latter. Even if that is a coherent possibility, it would be the case that "water" refers to both H<sub>2</sub>O and ghost water and *not* that water refers to ghost water alone. So even in that possibility it wouldn't be the case that H<sub>2</sub>O is merely nomologically connected to water.

<sup>20</sup> If the phenomenal state and the correlated physical state are the same then the correlation holds in all metaphysically possible worlds but if dualism is true the correlation may hold only in worlds in which there are laws linking the distinct phenomenal and physical states. The term "correlation" is neutral between metaphysical and nomological connection. If a correlation is found, i.e., if we have inductive evidence that N or F is coextensive with P in humans, the physicalist will posit the relationship to be metaphysical, while the dualist will posit a nomological connection. In either case, these correlations are supposed to be exceptionless, i.e., non-statistical in nature.

We will talk more about the explanatory gap in Section 2. Let's now focus on the search for the neural correlate of phenomenality. There is an underlying assumption that guides the scientific study of the mind, namely that phenomenal states will turn out to be – or to be nomologically connected to – genuine, unitary physical (neurophysiological) or functional states<sup>21</sup>. I have been presupposing, along with most philosophers who address these issues, that it is clear exactly what a physical state is and that not every predicate that is physically definable – e.g., by disjoining physical predicates – denotes a genuine unitary physical state. Similarly, I have been presupposing that not every specification of a causal role or disjunction of causal roles characterizes a genuine functional state. The usual thought is that genuine states (functional, physical or otherwise) in some way “cut nature at the joints”; that they are the states that figure in laws of nature.

Though at the moment it is still an open question if this assumption will turn out to be correct, the alternative, i.e., that phenomenal states are not correlated with any unitary neurophysiological or functional states in all creatures that share the state but rather with a disjunction of physical or functional states, seems to physicalist phenomenal realists like Loar (1997) tantamount to eliminativism unless one posits a genuine non-physical state that unifies all these instances. It is an interesting question whether this is correct, but, to simplify matters, I would like to put the whole issue aside and I will assume that the search for a neural correlate of phenomenality is a viable project.

### *The Harder Problem*<sup>22</sup>

There is a further issue that complicates the physicalist position which goes beyond the Hard Problem. At first, it seems that once the neural correlates of phenomenality are found, physicalism will be in a good position to identify phenomenally conscious organisms and the kinds of

---

<sup>21</sup> It is usually (but not universally) thought that the functional states that are candidates for identification with phenomenal states lie somewhere between the neurophysiological correlates of phenomenal states and the superficial functional roles associated with them cashed out in terms of stimuli and behavior associated; i.e., they are thought to be some “deep” functional roles that our brain states have in terms of information processing.

<sup>22</sup> I am not following very closely Block's exposition of the Harder Problem in Block 2002. Rather, I set up the Harder Problem in a way that helps connect it to what I call the Hardest Problem.



phenomenal state they have – since if physicalism is true then phenomenal facts are objective physical or functional facts and so we should be able to tell whether they obtain (and whether our usual methods of attributing them on the basis of behavioral evidence are accurate). But even if we assume, as I will for the moment, that this can be done with respect to other humans, the physicalist may still not be in a position to determine whether other beings – dogs, fish, Martians, robots, etc. – are phenomenally conscious or what the character of their phenomenal states are. The trouble is that even if the neural correlates of phenomenal state P are identified in humans, there are systematic reasons why there would be more than one candidate state – N, F<sub>1</sub>, F<sub>2</sub>, etc. – to play this role and so why we would still not be in a position to tell whether P is instantiated in other beings.

Here is why. Canonical evidence for the neural correlates of phenomenality comes from the human case. But quite plausibly, for every neurophysiological state that is nomologically correlated with a phenomenal state in humans, there are one or more functional states that are also nomologically correlated with that phenomenal state in humans. And since phenomenal concepts refer to phenomenal states directly – and not through a neurophysiological or functional definition – reflection on them is not sufficient to identify which of these is the actual referent. This fact is related to the Hard Problem: because physicalists don't have an explanatory metaphysics of phenomenality, even if neural correlates for every single phenomenal state in humans are found, they still won't be in a position to identify phenomenal states with neuro-physiological as opposed functional states or some mix of them. That means that physicalists also won't know which phenomenal states are to be projected beyond ourselves, to animals, Martians, robots, and so on. If the search for the neural correlate of phenomenality in humans proves fruitful, brain research can narrow down the range of states that might be identified with phenomenality considerably. But in all likelihood, it won't be narrowed down enough to settle on a single candidate. Though this epistemic handicap is not unique to physicalists – dualists are also not in a position to know the distribution of phenomenal states over other species – this ignorance is especially problematic for physicalists who claim that the distribution of physical properties fixes the distribution of all properties.

You might wonder why it should be the case that phenomenal states are nomologically correlated in humans with a range of neurophysiological and functional states. It may turn out that we have inductive evidence to the claim that P is nomologically correlated in humans with N. The problem is

this: it follows *with logical necessity* if N (and so P) is instantiated, that there are multiple functional states  $F_1, F_2 \dots F_n$  (one of whose realization is N) that are also instantiated. And while it is possible in principle that some (or all) of these functional states are *not* nomologically correlated with P in humans – that, e.g.,  $F_i$  can be instantiated in the absence of N (and so P) – it seems very plausible that, as a matter of fact, *at least some* of these functional states *are* nomologically correlated with N (and so P), i.e., that they don't have nomologically possible realizations in humans other than N. If that is the case, P has *multiple correlates* in humans. Though this is ultimately an empirical matter, I find the thesis of multiple correlates very plausible. I also find it plausible that some of these functional states are shared by other species even when these species don't share with us their neurophysiological realization.

This means that as far as we know – assuming physicalism, and the existence of these correlations – phenomenal states might be neurophysiological, or they might be functional. Because of the explanatory gap, N and F always seem compatible with the presence or absence of the relevant phenomenal state. If this is correct, then it is impossible to know what (if any) phenomenal states might be realized in other creatures that share with us the relevant functional organization but not the particular neurobiology or physical make-up.

This is not quite the skeptical problem of Other Minds (though it is related to it). Even if you dismiss skepticism about whether other humans are phenomenally conscious – after all, they share with you all of the possible candidates for neural correlates of phenomenality – you are still in no position to know about the phenomenal states of other creatures, and you don't even have a conception of how to remedy this situation. Even if you knew, on the basis of inductive evidence, all the neural correlates of phenomenality, you still wouldn't know what it's like for a cat, a fish, a Martian or robot to taste a lemon. There is a sinking feeling that this is wrong. That if physicalism is true then someone knowledgeable about the physical facts should be knowledgeable about all the facts. A satisfactory physicalist account, at a minimum, should explain why this is not so.

*The Hardest Problem*

The previous discussion presupposed that phenomenal concepts, such as the concept REDISH<sup>23</sup> refer determinately to objective states – modulo the vagueness exhibited by most general concepts. It is presupposed in our understanding of the Harder Problem that, though there are indeterminate cases due to vagueness, e.g., it might be indeterminate whether a chimpanzee on a given occasion is having a reddish experience or a bluish experience, often there *is* a determinate fact of the matter about experience, in other words, it very well *might be* a determinate fact on some other occasion that they are undergoing a reddish experience. We hold it as a central piece of folk wisdom, and I have presupposed it in my discussion of the Hard and Harder Problems, that there is *generally* a matter of fact – even if, as the Harder Problem indicates, we are incurably ignorant about it – about whether this or that creature has this or that phenomenal state. But further reflection on the Harder Problem casts this assumption into doubt. This is the Hardest Problem. The Hardest Problem for physicalism is a problem of reference: how do phenomenal concepts that, in their canonical, first person, introspective applications, refer directly to a thinker’s own phenomenal states succeed in referring *determinately* to objective states of this thinker?

The worry arises from the same issue we have discussed before: it is plausible that there are multiple distinct – neurophysiological and/or functional – correlates of phenomenality in humans. The epistemological consequence of this is the Harder Problem, that is, the problem that we cannot know if other species or even sophisticated robots share our phenomenal states. But the Harder Problem was predicated on the assumption that, modulo vagueness, there generally *is* a matter of fact about these questions. In the usual case, a member of another species, or a sophisticated robot, either does or doesn’t share our phenomenal states. The Hardest Problem is that it is hard to see how this can be true, given physicalism. In other words, it is hard to see how phenomenal concepts can refer determinately to one of these co-instantiated brain states to the exclusion of others.<sup>24</sup> What, if anything, determines which of these neurophysiological or functional states are the

---

<sup>23</sup> The concept REDISH refers to a *kind of phenomenology* – as opposed to the color red itself – normally had by looking at red colored objects.

<sup>24</sup> See Papineau 2003 for a discussion of all three problems. Papineau in the end embraces the view – which he thinks is a consequence of physicalism – that phenomenal concepts don’t have determinate reference and consequently that there is no matter of fact about application of phenomenal concepts to creatures physically different but functionally similar to us. In Section 3 I will take up whether this view entails eliminativism about phenomenality.

referents of phenomenal concepts?<sup>25</sup>

The “phenomenal concept strategy”, first deployed by Loar (1990, 1997) has been shown to provide a plausible physicalist solution of the Hard and Harder Problems.<sup>26</sup> I propose that it also holds the key to the Hardest Problem. Let us now look at the general account of phenomenal concepts that underlies the phenomenal concept strategy.

## 2. Three problems in light of an account of phenomenal concepts

One of the ways physicalists have tried to deal with the Hard and Harder Problem is to deny the special status of phenomenality. This route is predicated on the belief that the conceivability of zombies and the existence of the explanatory gap are not compatible with physicalism. It has been proposed that phenomenal concepts can be analyzed in non-phenomenal terms, e.g., in terms of functional role<sup>27</sup> or representation<sup>28</sup>. If these analyses were successful, the explanatory gap could be bridged, and the conceivability of zombies denied. All we would need to do is find the neurophysiological states that realize the functional or representational states figuring in the analysis of phenomenal concepts.<sup>29</sup> I find this position very unintuitive. I think that phenomenal concepts refer to phenomenal states directly and they are not functionally analyzable.

Others have tried to close the gap by proposing scientific accounts of phenomenality that are explanatory in the way the genetic theory of the transmission of traits is explanatory. It has been suggested, e.g., that being a phenomenal state is being processed in a global workspace;<sup>30</sup> that

---

<sup>25</sup> The Hardest Problem doesn't arise for dualism – on that view, phenomenal states are unitary non-physical states, and they are the unique referents of phenomenal concepts.

<sup>26</sup> For further developments of the strategy, see, for example, Papineau (2002), Block, (2002, 2006), McLaughlin (2006), Balog (2012b).

<sup>27</sup> Lewis 1966, e.g., is a clear example.

<sup>28</sup> See, e.g., Jackson 2003.

<sup>29</sup> This latter is predicated on the assumption that representational states can in turn be analyzed in terms of functional/causal relations.

<sup>30</sup> Baars 1988.

phenomenal states are representational,<sup>31</sup> that they involve higher order thought;<sup>32</sup> that there are "dynamic sensory-motor contingencies" associated with each sensory modality and these sensory-motor contingencies are somehow explanatory of the qualitative character of the particular sensory modality;<sup>33</sup> and that a state is phenomenal if it encodes highly integrated information.<sup>34</sup> In my view these suggestions, too, miss the mark. These theories are no more explanatory than neurophysiological accounts. The question still remains unanswered: why couldn't it be that different (or no) phenomenal states arise, for example, from those particular sensory motor contingencies? None of these theories can give an explanatorily satisfying answer for this question.

There are also those who, having lost hope in a physicalist solution to the Hard Problem, deny that phenomenality is real.<sup>35</sup> On the other hand, those physicalist phenomenal realists who think that physicalism is compatible with the conceivability of zombies and the explanatory gap (Hard Problem), and the lack of epistemic access to the phenomenal states of creatures with different neurophysiology (Harder Problem) have to come up with a convincing story about why this is so – how it is that phenomenality is physical yet it is impossible to produce an explanatorily satisfying theory about its scientific nature that would enable us to detect if snails, or any other creatures, are phenomenally conscious, and would reveal the illusoriness of the zombie-scenarios right from the start. This story cannot evoke the idea that phenomenality is metaphysically special. After all, a phenomenal state, on this view, is just another physical or functional state, 40 Hz oscillations, or being processed in a global workspace or what have you. The most promising story – the only promising story, in my view – has to do with the special way phenomenal concepts operate. I suggest that there already exist the broad outlines of such an account.

It can be shown that this account not only provides a physicalist resolution to the Hard and the

---

<sup>31</sup> See, e.g., Harman 1990, Tye 2000.

<sup>32</sup> See Rosenthal 2002.

<sup>33</sup> See, e.g., Hurley and Noë 2003.

<sup>34</sup> Tononi et al. (2016).

<sup>35</sup> See, e.g., Rey (1988), Dennett (1991), Frankish (2016).

Harder Problem but also explains the other perplexing epistemic features of phenomenality as well.<sup>36</sup> Here, however, I am going to focus exclusively on those three problems.

*The constitutional account of phenomenal concepts*<sup>37</sup>

The key idea underlying the phenomenal concept strategy is to give a physicalistic account of how phenomenal concepts can refer to phenomenal states *directly* and yet in a *substantive* manner. To spell this out I need to cover some background on the nature of concepts in general.

I take concepts to be constituents of thoughts and I take concepts and thoughts to be mental representations. I will also assume that concepts and thoughts are language like<sup>38</sup> – concepts are the words of *Mentalese*.<sup>39</sup> The important point for the following is that since tokening concepts and having experiences are occurrent entities (events, processes) they can be constituents of one another or bear causal relations to one another.<sup>40</sup>

Tokens of a concept, e.g., DOG possess a number of different properties and relations that are relevant to this discussion: i) realization properties, ii) conceptual role, and iii) semantic properties.

i) When one tokens an instance of DOG, say in thinking the thought DOGS BARK, that token is realized by some neural property. The neural properties that are relevant to the token's being a token of DOG are its *realization properties*. A concept's realization properties are analogous to the physical

---

<sup>36</sup> For a discussion see Balog 2012a.

<sup>37</sup> This section loosely follows my discussion of the same topic in Balog 2012a.

<sup>38</sup> There are non-language like, non-conceptual mental representations – image-like, map-like representations – as well. It is plausible that states with phenomenal properties, such as perceptions and sensations, are such representations.

<sup>39</sup> This account of thought was originally proposed by Fodor (1975).

<sup>40</sup> There are philosophers who would like to avoid Mentalese or avoid positing representations altogether. It may be that my account can be made compatible with their ontologies but that is not something that I can do here.

properties that realize a spoken, written, or electronic token of “dog”.

ii) A concept’s *conceptual role* is the totality of causal relations (and dispositions) that tokens of thoughts containing the concept bear to each other and to perceptual inputs and behavioral outputs. Certain aspects of a concept’s conceptual role may be essential to or even individuating of that concept while others are merely accidental; e.g., it is essential to the concept OR that one be inclined to make certain inferences, such as the inference from  $P$  to  $P \vee Q$ . It might also be essential to perceptual concepts, e.g., RED, that they be caused by certain perceptual inputs. Presumably, however, it is not essential to RED that one be caused to believe RED IS MY FAVORITE COLOR by the same perceptual inputs. How exactly to draw the distinction (which may be vague) between a concept’s essential and non-essential roles is controversial.<sup>41</sup>

iii) A concept’s *semantic properties* concern what, if anything, the concept refers to. For example, the concept DOG, refers to the property of being a dog. Exactly what determines the reference of a Mentalese word (with particular realization and role states) is a difficult and controversial matter.<sup>42</sup> It is widely (though not universally) held that a concept’s role (or the part of it that is essential to the concept) at least plays a part in determining the concept’s reference. This part is the concept’s *mode of presentation*. It often, but not always, has the form of a description - i.e., the thinker is disposed to infer the description from the tokening of the concept - i.e., from ARISTOTLE one is disposed to infer THE TEACHER OF ALEXANDER, etc. One can think of these descriptions as contents of a file attached to the concept. It is also widely accepted that reference is determined at least partly by external – causal, informational, or teleosemantic – relations of the concept to its environment.

A thinker typically has only partial epistemic access to features i-iii by introspection. When I attend to my thoughts, I typically can obtain introspective knowledge of their semantic contents, e.g., that I am thinking about dogs. It is also plausible, though controversial that one can obtain information about the conceptual roles of one’s concepts – and which of these are essential – by intuitions based on thought experiments, e.g., by asking oneself questions like “could one know  $p$  if  $p$  were false?”.

---

<sup>41</sup> This is as difficult a problem as the others I am discussing but there is no space here to discuss it.

<sup>42</sup> Another hard problem not to be addressed here.

But the realization properties of one's Mentalese words – the “shapes”, or “mental ink” they are written in, so to speak – are almost always completely opaque. Almost always, with the exception – I propose – of *phenomenal concepts*.

I propose an approach to phenomenal concepts that fits into this general framework but that also showcases the special nature of phenomenal concepts. An examination of phenomenal concepts suggests that a successful account of them will posit an intimate connection between phenomenal states and the concepts we form of them. Loar (1990, 1997) suggested the idea that phenomenal concepts are *direct recognitional* concepts. Abstracting from some of the details, the core idea that I take him to express is that when a person is having a particular experience, they can deploy a concept that refers directly to the experience in that the referent is somehow present in the concept's *mode of presentation*. How is one to understand this idea?

There is an ambiguity in Loar's account that points the way toward an answer. Loar sometimes describes phenomenal concepts as in some way “tracking” their referents. This indicates that – despite his suggestion that the mode of presentation in some way involves the referent itself - he is thinking of a phenomenal concept and its referent as distinct entities related by causation. But such an account leaves too much of a distance between a phenomenal concept and its referent; it allows, for example, for the seemingly absurd possibility that when the causal mechanism connecting phenomenal states and phenomenal concepts breaks down you might introspect having a certain kind of phenomenal experience when in reality you are having none of the sort. Causation seems to be the wrong model for how the reference of phenomenal concepts is determined.

The solution is to posit a relationship between phenomenal concepts and phenomenal states that is more intimate than causation: to hold that each canonical (first person, introspective) tokening of a phenomenal concept is constituted by a phenomenal experience whose phenomenal properties the concept refers to.<sup>43</sup> Imagine seeing a red flower and wondering whether others have the same kind of phenomenology as you do. You think to yourself: I WONDER IF OTHERS' EXPERIENCE IS ALSO REDDISH WHEN SEEING THIS FLOWER. Your tokening of the concept REDDISH

---

<sup>43</sup> Similar ideas are proposed in Papineau 2002, David Chalmers 2003, Block 2006, and Balog 2012a.



here is constituted by the particular instance of reddish phenomenology that the flower occasions. That is the “mental ink” in which this instance of REDDISH is written. In other words, according to this view, phenomenal concepts (in their canonical applications) have very special *realization properties*: the neural properties realizing them are instances of the very same neural properties they refer to. But what is so special about phenomenal concepts, on this account, is not only that their *realization properties* are instances of their *referents*, but that this very fact is crucially involved in determining their *semantic properties*. Not only are these concepts realized by instances of their referents, but they refer to what they do at least in part *in virtue* of this fact. Their realization properties and conceptual role together determines their referent. This is, of course, very different from any other concept. Most concepts are not realized by tokens of their referents at all; but even those, like the concept ATOM, that are, mean what they do completely independently of this fact about their realization.

There are many details about this account that need to be worked out.<sup>44</sup> What I propose here is to see what light this account can shed – assuming all those details can be hammered out – on the three problems of phenomenality we have been talking about.

a) *The conceivability of zombies* is explained by the *directness* of our phenomenal concepts. Phenomenal concepts are supposed to be different in this way from concepts like WATER and even name concepts like CICERO. Chalmers and Jackson (2001) claim that these concepts are associated *a priori* with descriptions (e.g. “the transparent potable liquid...”, “the Roman orator who is at the origin of a causal chain culminating in this token”) and these connections are *sufficient* to rule out a priori a scenario where, e.g., everything is physically the same but yet there is no water. One doesn’t have to commit to this to see that zombies are conceivable; however, the conceivability of zombies is only really significant if this is the case. So the point is that even if one allows this to be true with respect to WATER, or CICERO, it is clearly not so with respect to phenomenal concepts; and *that* is the reason why the existence of zombies cannot be ruled out a priori.<sup>45</sup> The directness of

---

<sup>44</sup> I have done some of this in Balog 2012a.

<sup>45</sup> As I pointed out earlier there are dissenters who think phenomenal concepts have descriptive analyses; see e.g., Lewis 1966, Jackson 2003. I will ignore this possibility and proceed on the assumption that, as Chalmers suggests, zombies are conceivable.

phenomenal concepts, on the constitutional account, is explained by the fact that their reference is determined by how they are constituted – and *not* by any description that is associated *a priori* with them – and, importantly, this fact is perfectly compatible with physicalism. One can hold the constitutional account, and therefore be comfortable with the conceivability of zombies, whether one is a physicalist or a dualist.<sup>46</sup> The conceivability of zombies is explained solely by the special cognitive architecture involved in phenomenal concepts which is orthogonal to the issue of the metaphysical nature of phenomenality.

b) *The explanatory gap*. Recall that the problem of the explanatory gap is that no amount of knowledge about the physical facts (brain functioning and so on) is able to explain why a particular brain state has a particular feel, e.g., feels giddy. This contrasts with the way the fact that water is composed of H<sub>2</sub>O molecules together with physical and chemical laws explains why water is potable, transparent and so on. Once we have an explanation of why H<sub>2</sub>O behaves in watery ways (and that it is the only substance that does so) we have an explanation of identity statement. The hypothesis that water=H<sub>2</sub>O is quite natural in the light of all the above information<sup>47</sup> – indeed, the opposite hypothesis strikes one as borderline incomprehensible. The idea that the processes involving H<sub>2</sub>O molecules are only nomologically correlated to the non-physical and non-chemical processes involving water is a non-starter.

On the other hand, the hypothesis that a phenomenal state is *not identical* with some neurophysiological/functional state of the brain is quite compatible with the evidence. Since we can't explain why a brain state feels giddy in neurophysiological terms, the dualist hypothesis that phenomenal states and brain states are merely nomologically correlated makes perfect sense.

The reason the explanatory gap cannot be closed in the case of phenomenality is that, while we do not have any special access to the nature of water that is not based on physical or functional

---

<sup>46</sup> Chalmers 2003, e.g., proposes a dualist version of the account.

<sup>47</sup> This is true even if one doesn't believe that the identity statement can be a priori derived from information about H<sub>2</sub>O and conceptual truths involving 'H<sub>2</sub>O' and 'water'.

information, we do have *substantial* insight into the nature of phenomenal experience that is not mediated by such information. As far as we know, the nature of phenomenality might elude any physical or functional understanding; it might be something over and above the physical.

However, the constitutional account can explain our *direct* and *substantial* grasp of phenomenal states – and so the explanatory gap – in a way that is perfectly compatible with physicalism. It explains it in terms of the constitution relation between concept and referent. When I form a phenomenal concept, I grasp its referent in terms of the shared phenomenality of concept and experience – *what it's like* to be in that state is the same as what it's like to entertain the concept. When, e.g., I focus on the reddishness of my experience of a flower – not on what it represents but on its qualitative character – “holding” the experience up for inspection in my mind, my conception of it contains that very experience. Having the experience and thinking about it then shares something very substantial, very significant: the phenomenal character of the experience. It is this shared phenomenality that produces the sense that one has a direct insight into the nature of the experience.

However, as is clear from what we have said, this kind of direct insight into the nature of phenomenal experience does not reveal anything about the metaphysical nature of phenomenality. It is very different from the kind of insight that a scientific analysis of the brain would provide into the nature of a brain state. The one involves *having* the state, the other, *analyzing* it into its components. Those are very different activities. On the constitutional account, the existence of the explanatory gap – a product of our substantial grasp of phenomenal states – does not preclude the possibility that phenomenality is purely physical.

*c) The Harder Problem.* Physicalists need to provide an account of why it is impossible to know if creatures physically different from us share phenomenal states with us. This problem is connected to the Hard Problem. If we understood how phenomenality arises from brain processes in the way we understand how water arises from a collection of H<sub>2</sub>O molecules, or digestion arises from chemical processes in the digestive system – that is, if we could close the explanatory gap for phenomenality – we would be able to tell, for any physical or functional state, whether it is a phenomenal state. We would also be in a position to know – in principle at least – what phenomenal states, if any, other

creatures have in common with us. But we can't.

How does the constitutional account help reconcile us to such ignorance? It does by giving an account why the explanatory gap arises that is still compatible with physicalism. Given the constitutional account of phenomenal concepts, such ignorance is what we should *expect*, even on a materialist assumption. We do not know what phenomenal states if any non-human individuals have because we know our own phenomenal states *qua* phenomenal in a very special way. It is the nature of these concepts, i.e., the fact that we directly apply them to phenomenal states that explains why no perspicuous physical explanation of phenomenal states can be found, and why, consequently, we can't know if creatures different from us share phenomenal states with us. There is no need to suppose that physicalism is false in order to explain why the explanatory gap arises. Physicalism has the resources to settle the questions posed by the Hard and Harder Problem.

### *The Hardest Problem*

The constitutional account of phenomenal concepts promises to resolve the conceptual and epistemic puzzles raised by the Hard and the Harder Problem – even if it doesn't quite dispel the illusion of mystery. The Hard and Harder Problem, in this telling of the story, persists as an illusion, a metaphysical version of the Müller-Lyer illusion. The constitutional account, however, runs into difficulty with the Hardest Problem.

In the previous discussion I have tacitly assumed phenomenal realism – according to which there is a matter of fact about what phenomenal qualities are instantiated, e.g., whether chimpanzees have the same phenomenology when looking at red objects as we do. There is a further, related intuition that it is a determinate matter of fact whether our phenomenal concepts apply to other creatures as well – modulo the familiar issue of vagueness and borderline cases – even if we can never find out. The Hardest Problem is that it is very hard to see, on a physicalist account, how this could be the case.

In what follows, I will argue for two claims. The first is that the Hardest problem, by presenting a serious challenge to the determinacy of phenomenal concepts, also threatens phenomenal realism,

and with that, our solution to the Hard and Harder Problems. The second is that, luckily for the physicalist, one needn't hold the indeterminacy of phenomenal concepts to uphold physicalism. In the end, the constitutional account of phenomenal concept can answer the questions raised by the Hardest Problem without resorting to indeterminism.

The problem when it comes to phenomenal concepts is that there is good reason to suppose that there are multiple neural and functional states that are nomologically coextensive with any given phenomenal state in humans. If this is the case, then it is very hard to see how there could be a matter of fact about which of these states our phenomenal concepts refer to, *especially* given the theory of phenomenal concepts we discussed above. To explain why, we will have to go into a little more detail about the constitutional account.

Direct phenomenal concepts pick out their referent from the subject's point of view. In this way, phenomenal concepts are like the indexicals 'I', 'now', 'here'. But while we understand perfectly well how 'here' picks out an objective, determinate spatial location (since the thinker has an objective spatial location at any time and 'here' picks it out), things are much less straightforward for phenomenal concepts. The constitutional account claims that, e.g., REDDISH picks out its referent via an instance of *that very property*, reddishness. We conceptualize reddishness via introspectively attending to an instance of it in our experience and thereby incorporating it in a phenomenal concept. But how does this cognitive machinery determine the proper reference of the concept thus based in an instance of the referent?

On the constitutional account, REDDISH can be understood as the concept \*this attended reddish experience\* where '\*' stands for whatever concept forming mental mechanism is involved in turning an experience currently attended by the subject into a concept that represents the very property it instantiates; and "this attended reddish experience" simply stands for a particular experience the subject is attending. There is a further complication which I have to put to the side for now, involving the conceptual machinery that determines which exact phenomenal property is being

referred to, given that a phenomenal concept's constitutive experience is an instance of many different properties, e.g., reddish, dark reddish, crimsonish, etc.<sup>48</sup>

The question is, how could a phenomenal concept C of a phenomenal property P (e.g., reddish) pick out a neural property N, to the exclusion of some functional property F that is equally nomologically coextensive with P in humans (though perhaps not in other organisms)? Suppose P is identical with F, but *not* with N. How could phenomenal concepts track *that* fact? It is hard to see how; phenomenal concepts do not seem to have the right kind of built in complexity to disambiguate between these different but coextensive properties.<sup>49</sup>

David Papineau (2003) argues, on the basis of these considerations, that phenomenal concepts refer indeterminately; and that consequently, that there is no matter of fact of whether other species, or certain kinds of mere functional replicas share our phenomenal states.

Attempting to make it sound like this is not a problem for physicalism, he summarizes the situation in the following way:

But I do not want to argue that there is something less than definite in the doppelganger's [a functional but not physical duplicate of a human being] experience itself. The doppelganger definitely feels as it does, however that is. My thesis is only that it is indeterminate whether it is *pain*, where I take this indeterminacy to arise from vagueness in our human term 'pain'. The indeterminacy doesn't lie in the doppelganger's experience itself, but in whether that experience is similar enough to cases of human pain to fall under our term 'pain', a term whose content derives from exemplars provided by human pain. (2003, p. 218)

There are a number of problems with this view.

---

<sup>48</sup> Plausibly, this has to do with *phenomenal similarity*. One idea is that, however phenomenal similarity itself is constituted, a phenomenal concept's extension is determined by an associated disposition to make the relevant similarity judgments, based on a comparator that, for example, uses a stored template or the like. So, my concept REDDISH refers to reddish in virtue of my disposition to deem reddish – as opposed to crimsonish – experiences as falling under it based on my recognition of their similarity to a stored template of a reddish experience. For more discussion of this, see Balog 2012a.

<sup>49</sup> Notice that dualism, whatever other unattractive consequences it might have, does not face the Hardest Problem in the same way. A dualist can merely posit phenomenal states to be the unambiguous referent of phenomenal concepts.

First, the supposed indeterminacy of phenomenal concepts is peculiar and is different from other familiar cases of referential indeterminacy. Take the case of the concept HEAT which was once used for both temperature (the average kinetic energy of the molecules in an object) and amount of heat (the total amount of kinetic energy of the molecules in an object). But while in this case progress in science pointed to an ambiguity in the concept which resulted in the formation of two distinct concepts replacing the old unitary concept HEAT (i.e., TEMPERATURE and AMOUNT OF HEAT), no such possibility exists for the phenomenal concept PAIN. On Papineau's account of how phenomenal concepts work, one could only refine the concept PAIN to refer determinately to either  $N_{\text{pain}}$  or  $F_{\text{pain}}$  (the neurophysiological and functional states that are each nomologically coextensive with pain in humans) in non-phenomenal ways, employing physical and functional descriptions to pick out their respective referents. By his own admission, no *phenomenal concept* could achieve such disambiguation.

The supposed indeterminacy of phenomenal concepts also differs from the indeterminacy of vagueness. There are colours that fall somewhere between, say, red and orange, so it is indeterminate which concept applies, but there are also colours that fall under RED determinately. Though phenomenal concepts do have this kind of indeterminacy, the indeterminacy Papineau talks about is not of this sort. According to him, there are no instances of REDDISH that apply determinately to  $N_{\text{red}}$  or  $F_{\text{red}}$ .

The indeterminacy Papineau has in mind is also different from the referential indeterminacy Quine suggested concepts such as RABBIT possess. RABBIT, according to Quine, refers indeterminately to rabbits, undetached rabbit-parts, momentary rabbit stages, etc. This referential indeterminacy, however, doesn't extend to propositions in that, e.g., THERE IS A RABBIT is true or false for all permissible assignments of reference. On the other hand, if REDDISH refers indeterminately to  $N_{\text{red}}$  and  $F_{\text{red}}$  then THIS MONKEY HAS A REDDISH EXPERIENCE might come out true under one interpretation ( $F_{\text{red}}$ ) but false under the other ( $N_{\text{red}}$ ).

More importantly, though, if there isn't something "less than definite in the doppelgänger's experience itself" then, it seems, there is something definite about my experience and its phenomenal character as well. As far as I can see, this is an affirmation of qualia realism – the view

that there is a determinate matter of fact about the phenomenal character of the experience of conscious subjects. However, Papineau is not entitled to hold both that our concept PAIN refers indeterminately to  $N_{\text{pain}}$  and  $F_{\text{pain}}$  and also that pain itself is a “definite” state about whose phenomenal character there is a matter of fact.

On the natural assumption that PAIN refers to pain, which is a “definite” state of the subject, PAIN cannot also refer to both  $N_{\text{pain}}$  and  $F_{\text{pain}}$  indeterminately since, unless they are themselves identical, only one of them can be identical with the “definite” property pain. It follows then that accepting the indeterminacy of phenomenal concepts is incompatible with phenomenal realism and so it leads to eliminativism about phenomenal states.

But it is hard to be an eliminativist about phenomenal states. The Hard and Harder Problem, the explanatory gap, and all the other mysteries of phenomenality stem from the fact that we not only refer to phenomenal states directly – without phenomenal or functional modes of presentation – but we refer to them in a *substantive* way. We are *acquainted* with the referent of phenomenal concepts and from the first person point of view of that acquaintance – as opposed to the third person point of view from which we try to understand how phenomenality fits into a “scientific image” of the world – it makes little sense to suppose that there are no “definite” phenomenal states.

A physicalist qualia realist then is committed to the view that PAIN *does* refer unambiguously to either  $N_{\text{pain}}$  or  $F_{\text{pain}}$ . I think such a position is actually quite plausible for the physicalist. To see why, we need to remind ourselves of how the Hard, Harder and Hardest Problem arises. They arise because we try – and fail – to understand phenomenality and its place in the physical world from the third person point of view. In the case of the Hardest Problem, we are led to the view of referential indeterminacy via our failed effort to understand how reference gets disambiguated from the third person point of view. Since nothing from that point of view reveals how  $N_{\text{pain}}$  but not  $F_{\text{pain}}$  could be constitutive of a phenomenal concept, indeterminacy seems to follow.

I now propose that the resolution of all three problems is the same as well. In each case, the constitutional account of phenomenal concepts helps us recognize the problem as an illusion created by the special cognitive architecture of the first person concepts we apply in introspection and their



isolation from third person conceptions. To put it another way, the constitutional account helps us see how we can be led astray by our predilection to assimilate the first person to the third person.

When we focus on the role of introspection in phenomenal concepts, we can see how, *on the assumption* that pain is either identical to  $N_{\text{pain}}$  or  $F_{\text{pain}}$ , phenomenal concepts could manage to pick out the correct referent, without that being discernible from the third person point of view. Since *only phenomenal states can be attended*, and therefore quoted in the mind in introspection – whereas non-phenomenal brain states cannot – it is evident how PAIN, a concept that works by quoting an instance of the referent, picks out whichever of  $N_{\text{pain}}$  and  $F_{\text{pain}}$  is identical to pain. If, for example, pain is  $N_{\text{pain}}$  then that is what is being mentally quoted – since non-phenomenal states such as  $F_{\text{pain}}$ , could not be present in the mind for introspection. The way phenomenal concepts single out they referent cannot be understood from the third person point of view for the same reason that the phenomenality of pain cannot be understood from the third person point of view; because of the radically different cognitive machinery involved in first person phenomenal and third person physical/functional concepts.

Of course, if physicalism is true, facts about the reference of phenomenal concepts, like everything else, is grounded in physical facts, in this case physical facts about the brain. However, considered from the third person, scientific perspective, these facts cannot explain either the identity of pain with, say  $N_{\text{pain}}$  rather than  $F_{\text{pain}}$ , nor the subjective nature of attention to phenomenal states, nor the nature of phenomenal concepts; nor, for that matter, why non-mental, non-phenomenal states cannot be attended. This is just a manifestation of the Explanatory Gap, which we have dealt with in the discussion of the Hard and Harder Problem. The Hardest Problem, exactly like the Hard and Harder Problem, turns out to be another artifact of an irreducible conceptual dualism with regard to phenomenal experience.

This account of referential determinacy in the face of the Hardest Problem invokes the spirit of Loar's (1995) attempt to account for referential determinacy in the face of Quinean considerations to the contrary. Loar argues that our first person concept of reference tracks which of various external relations qualify as reference for our concepts (so, that the concept THAT RABBIT indeed refers to that rabbit and not, say that rabbit surface), even while it is not possible to understand how

this can be from the objective, third person perspective. He argues for the radical thesis that the reason we cannot disambiguate reference from the third person point of view is that, understood properly, reference is essentially a subjective concept with a phenomenal component that – just like phenomenal experience itself – is not explicable from the objective perspective.

While my view doesn't presuppose Loar's controversial thesis about the concept of reference, my solution to the Hardest Problem is in the spirit of his crucial insight: that it is compatible with physicalism that phenomenal experience make sense – in a certain way – only from the subject's point of view.

### References:

- Baars, Bernard J, 1988. *A cognitive theory of consciousness*. Cambridge University Press, New York.
- Balog, Katalin, 1999. "Conceivability, Possibility, and the Mind-Body Problem", *The Philosophical Review*, Vol 108, No 4 (October 1999). Reprinted in Volume XXIII of *The Philosopher's Annual*.
- Balog, Katalin. 2005. "Ontological Novelty, Emergence, and the Mind-Body Problem", in (Günter Abel, ed.) *Kreativität*, Hamburg: Meiner Verlag.
- Balog, Katalin, 2009. "Phenomenal Concepts". In Brian McLaughlin, Ansgar Beckermann, and Sven Walter (Eds.), *The Oxford Handbook of Philosophy of Mind* (pp. 292-312). Oxford: Oxford UP.
- Balog, Katalin, 2012a. "Acquaintance and the Mind-Body Problem". In Christopher Hill and Simone Gozzano (Eds.), *The Mental, the Physical* (pp. 16-43). Cambridge: Cambridge University Press.
- Balog, Katalin 2012b. "In Defense of the Phenomenal Concept Strategy". *Philosophy and Phenomenological Research*, 84(1), 1-23.
- Block, Ned, 2002. "The harder problem of consciousness", *The Journal of Philosophy*, Vol. XCIX, No 8, pp. 391-425.
- Block, N. 2006. Max Black's objection to mind-body identity. in *Oxford Studies in Metaphysics*, II, edited by Dean Zimmerman, Oxford University Press, 3-78.

- Braddon-Mitchell, D. and Jackson, F., 1996. *The Philosophy of Mind and Cognition*. Blackwell.
- Burge, Tyler, 1988. "Individualism and Self-Knowledge." *Journal of Philosophy* 85.
- Chalmers, David, 1996. *The Conscious Mind*, Oxford UP.
- Jackson, F. and Chalmers, D. 2001. "Conceptual Analysis and Reductive Explanation", *The Philosophical Review*, 110:315-61.
- Chalmers, David, 2002. "Consciousness and its Place in Nature", in (S. Stich and F. Warfield, eds) *Blackwell Guide to the Philosophy of Mind*, Blackwell.
- Chalmers, David, 2003. "The Content and Epistemology of Phenomenal Belief", in Q. Smith and A. Jokic (Eds), *Consciousness: New Philosophical Perspectives*, Oxford.
- Chalmers, D., 2009. "The Two-Dimensional Argument Against Materialism", in B. McLaughlin (ed.) *The Oxford Handbook of the Philosophy of Mind*.
- Chalmers, D, 2017. Panpsychism and panprotopsyism, in *Panpsychism: Contemporary Perspectives*, Oxford University Press.
- Dennett, Daniel, 1991. *Consciousness Explained*, Little, Brown and Company.
- Fodor, Jerry A. (1975). *The Language of Thought*. Harvard University Press.
- Fodor, Jerry, 1998. *Concepts: where cognitive science went wrong*, Oxford; New York: Clarendon Press.
- Frankish, K. (2016). "Illusionism as a Theory of Consciousness, *Journal of Consciousness Studies* 23(11-12).
- Harman, Gilbert 1990. "The intrinsic quality of experience", *Phil Perspectives* 4, Action Theory and Philosophy of Mind, pp. 31-52.
- Hurley, Susan and Noë, Alva, 2003, "Neural Plasticity and Consciousness", *Biology and Philosophy* **18**: 131-168.
- Jackson, Frank, 1982. "Epiphenomenal Qualia", *Philosophical Quarterly* 32, pp. 127-36.
- Jackson, F. 2003. "Mind and illusion". In (A. O'Hear, ed) *Minds and Persons*. Cambridge University Press.
- James, Williams, 1950. *The Principles of Psychology*, New York: Dover Publications.
- Levine, Joseph, 2001. *Purple Haze*, Oxford UP.
- Lewis, David, 1966. "An argument for the Identity Theory", *Journal of Philosophy*, LXIII. 1 (January 6,

1966): 17-25.

Loar, Brian 1990. "Phenomenal states", *Philosophical Perspectives* 4, Action Theory and Philosophy of Mind, pp. 81-108.

Loar, Brian, 1995. 'Reference from the First-Person Perspective', *Philosophical Issues: Content*, 6: 53-72. Reprinted in *Consciousness and Meaning: Selected Essays by Brian Loar* (eds. Katalin Balog and Stephanie Beardman), Oxford University Press, 2017.

Loar, Brian, 1997. "Phenomenal states", in *The Nature of Consciousness*, (Block, Flanagan, Güzeldere, eds.), MIT Press, pp. 597-616 (revised version of Loar 1990). Reprinted in *Consciousness and Meaning: Selected Essays by Brian Loar* (eds. Katalin Balog and Stephanie Beardman), Oxford University Press, 2017.

Lockwood, Michael, 1998. "The Enigma of Sentience", in Hameroff, S.R. et al, *Toward a Science of Consciousness II*, Cambridge, MA: MIT Press.

Maudlin, Tim, 1989. "Computation and consciousness", *The Journal of Philosophy* 86:407-432.

McLaughlin, Brian, 2006. "Identity, Explanation, and Consciousness: A Reply to Kim", forthcoming.

Nagel, Thomas, 1979. *Mortal Questions*, Cambridge UP.

Papineau, David, 2002. *Thinking about consciousness*, Oxford UP.

Papineau, David, 2003. "Could there be a science of consciousness?", *Philosophical Issues*, 13(1):205-220.

Rey, Georges, 1988. "A Question about Consciousness", in *Perspectives on Mind*, ed. H. Otto and J. Tueidio. Dordrecht: Kluwer.

Rosenthal, David, 2002. "Explaining Consciousness". In (Chalmers, ed.) *Philosophy of Mind*, Oxford University Press.

Stoljar, Daniel, 2001. "Two Conceptions of the Physical", *Philosophy and Phenomenological Research* 62: 253-281.

Stoljar, Daniel, 2005. "Physicalism and phenomenal concepts", *Mind and Language*, Vol. 20, Issue 5, 469-495.

Tononi, G; Boly, M; Massimini, M; Koch, C. 2016. "Integrated information theory: from consciousness to its physical substrate", *Nature Reviews Neuroscience*, 17: 450-61.

Galen Strawson, 2003. "Realistic materialism", In Louise M. Antony & Norbert Hornstein (eds.), *Chomsky and His Critics*. Blackwell.

Strawson, G., 2006. Realistic monism: why physicalism entails panpsychism, in *Consciousness and Its Place in Nature: Does Physicalism Entail Panpsychism?*, A. Freeman (ed.), Exeter: Imprint Academic.

Tye, Michael, 2000. *Consciousness, Color and Content*, ch. 3., MIT Press.