

The Impossibly Hard Problem of Consciousness:
Showing the hard problem of consciousness cannot be solved, and the way forward

Jorrit Jan Walinga

Abstract

I show the sort of solution currently commonly imagined as fitting the hard problem of consciousness is impossible to reach. The category mistake implicit in the problem can be dealt with, but the reformulated version will still never reach complete predictive ability and intuitiveness. Applications of this proof of impossibility and attempts by others to solve the hard problem are discussed. An overarching phenomenological-cognitive framework is proposed, partially to show the possibilities which are still open for those interested in discovering 'consciousness-processes'.

Introduction

Inquiry into the workings of the mind has caused one metaphysical problem to perpetually keep surfacing and has great thinkers struggling since as far back as the 17th century, namely the problem of the apparent duality of body and mind. Suitably, this problem is often referred to as the mind-body problem. It is the question of how the mind and the body are related. Yet another name for roughly the same problem, although already with the assumption of the physicalist worldview, is Levine's 'explanatory gap' (Levine, 1983), which is the gap between explanation of the physical and of the mental. But the contemporarily popular derivative of this problem which I will treat in this paper is known as the 'hard problem of consciousness' (which will be referred to with 'HPC'), using the definition from when it was first coined (Chalmers, D. 1995).

I will show how the HPC cannot get the sort of solution the problem implies, and how it is still possible to get an overarching framework for the easy problems.

I will work from a view which could be described as physicalist and 'process-dualist'. This implies that I do not assume anything supernatural interacting with, being produced by or producing the material world. This also applies to the mind, which is treated as a process emerging from the physical.

So consciousness will not be considered as something excreted by the brain, neither as a property of the neural substrates. Rather, I will treat it as the perceived 'inside' of a process through time. Explanations involving subjective experience and the mechanics of the brain are treated as first-person perspective and third-person perspective descriptions respectively. I consider this view to be shared by or, for the arguments of the paper, sufficiently coherent with mainstream thought to such a degree that I will leave out its defense.

Combining this metaphysical foundation with the most recent well-predicting scientific abstractions of the substrate 'underneath' our experience (the brain) leads us to the conclusion that all aspects of our phenomenology are changeable, without an external objective standard and evaluated internally. This results in a perspectivist (and internalist) stance as well, towards the entirety of our phenomenology.¹

A further, full explanation is beyond the scope of this paper, so this worldview can be treated as the ground of assumptions from which we build. However, these assumptions

¹ In this framework, metaphysics is seen as just a subsection of our model of the world. It is the case that our descriptions of our phenomena have underlying metaphysical assumptions, like usually a shared external outside world beyond the senses, but these are epistemologically dependent on the 'physics'(/phenomena) they underlie as 'doing metaphysics' is not necessary for epistemological processes to be seen as useful/truthful and what sort of metaphysics we have depends on our experiences.

do not have to be shared between me and the reader for the statements I make to be interpreted as internally valid, as the regular non-perspectivist framework can usually be translated into mine without problem, from the perspective of the perspectivist, in the following manner:

All claims stated as if they are objectively true are made knowing it is a useful shorthand for stating the claim is felt as true for me and that I believe the reasons for this feeling to be able to be roughly intersubjectively shared to such a degree that it is useful to publicly make the statement in the first place.

This paper will focus on how explanations are formed, when they feel sufficient and other epistemological concepts. So for the sake of coherence it is important I state in this introduction explicitly the implicit reasoning steps I will make later on.

The hard problem

'The hard problem', was first described by Chalmers as the question of why physical processes should give rise to subjective qualitative experience ("if there is something it is like to be that organism") to begin with. (Chalmers, D. 1995)

A description of what a solution, in line with the earlier philosophical framework, must look like is as follows: We have theories, systems of concepts, which map onto phenomena and model what we believe is happening in the world beyond our senses. Some parts of these models are about what we are from a scientifically reductive perspective, the brain, and we could hypothetically make models complex enough to perfectly predict large parts of the brain, the entire brain or even the brain with its immediate environment. (Which we could do by offloading some computation and data-storage in the world itself by manipulating the world in such a way that, when we interact with this artificial model, the same sort of functionalist systems of concepts activate in our minds as a close-up, slowed down, abstracted, etc. direct view of the brain would've produced if we had access to such a view. (i.e. writing things down and making computer models))

Another mode of looking at the world, or rather our world, is not imagining a causal chain of objects existing outside of ourselves responsible for the phenomenological changes but noting purely the patterns in the changes themselves. There is an important difference between having an experience and thinking about experience, although thinking about an experience is an experience as well. We need to have experiences to be able to think about experience, but once we've had them we can use the memories to abstract, generalize and project experience (in that order) to other minds including our former or future self.

The solution the HPC seeks is the logical steps, an explanation, which causes us to attribute experience, accurately, to (some system of concepts describing) physical processes. In the same way we can ask "Why does rain fall from clouds?" and accept an

explanation describing vapor condensation, the HPC asks how or why the brain produces consciousness.

Problematizing the problem

Brain processes, or whatever it is in which consciousness occurs, are asked by the HPC to be linked in such a way to experience that we can see them as two sides of the same coin at an intuitive level (for all cases, so an explanation of the fundamental ‘trick of consciousness’). This creates the following problems: What level of detail does the description need to have, does the necessary arbitrariness of that description matter and in what way? Is having such an intuition sufficient for being certain something is conscious? Can we imagine a process complex enough to be conscious in any recognizable way? Do we have any other way of validating whether something belongs to the class of conscious processes or not, other than through our own generalized experience?

Imagine, as some potential best-case scenario, we have some hypothetical supermind which has the computational ability to imagine all the causal steps happening inside a subsection of itself, from sensory input to abstract concept, of some part of experience. When shown a detailed recording of their brain they can tell you what sort of experience they were having at that moment. Their mental model of their brain is perfectly predictive of their experience and vice versa, but only locally to prevent a complexity paradox. Has this supermind solved the HPC? It still only correlates some type of brain process with some type of experience. The supermind can explain to you “when X occurs in my brain then I am conscious”, but it does not explain why X is a conscious process. The why-question informs about some causal connection, not a correlation. When we ask a why-question we want a reason/cause for that which the question is about.

Another thought-experiment: Suppose I show you someone being hit with a hammer. You can imagine that it hurts. The cause of this feeling is physical. I could even show you your own nerve-cells, or the brain processes connected to them, firing in response to being hit and over time you would understand that the firing of the nerve cells hurt on the same instinctual level as the hammer. But even though some physical causal chain is understood to be the cause of a certain experience it does not solve the hard problem because just connecting the physical and mental intuitively isn’t enough either. No ‘trick of consciousness’ was explained. And since the first instinct of ascribing the feeling of pain as a consequence to being hit with the hammer was a generalization and projection based on your own experience you could’ve easily failed to predict the right sort of feeling as well, if an important detail of the person being hit was different like being in a sedated state. Even if the causal links are broken down so precisely that we can predict perfectly when an event grouped under the concept ‘pain’ occurs, this model only predicts human experience and everything which is close enough such that we feel comfortable generalizing our model. It cannot tell us if a computer simulating a

brain in a typical Von Neumann architecture can feel pain (so more broadly: if functionalism is true).

Would combining the two thought experiments help? Or what if we make a supermind which automatically feeds an imagined experience into the rest of the brain such that thinking about a certain physical causal chain in the brain responsible for an experience produces that experience (and vice versa)? If thinking about certain brain processes which do not yet exist in the supermind would create them it would even solve the “Mary’s room thought experiment” (having full knowledge of an experience but not having had the experience itself, can you really know what it’s like to have that experience?). Such a mind would effortlessly, accurately and intuitively jump over the gap between third-person and first-person perspective, but it cannot say why conscious processes exist.

This why-question, by being formulated as a ‘why question’, is not just asking for a connection which links brain-processes to mental processes (which would be a ‘how’-question), but why that connection exists in the first place and not something else. It seems to be roughly equivalent to asking why subatomic particles (or whatever other mental model works in prediction) have the sort of behavior they do, and not something else. Breaking these subatomic particles up in smaller causal chains does not help, nor does more accurate prediction. If you can explain particles in terms of spacetime, and vice versa, you can still ask why that is the case. This continual questioning process is a property of the human mind which sometimes just does not want to be satisfied. Theistic religions use a god as the final cause, potentially as a stop to this questioning, and children seem to be particularly disposed to ask “why” questions presumably to expand their knowledge. But usually we are satisfied with some explanation. Why is this so for regular questions and is the HPC an exception?

An explanation logically connects one phenomenon to another, both abstracted into (systems of) concepts. “Why did the cup fall?” is satisfyingly answered by “The cat pushed the cup off the table.” We can imagine, because of our general knowledge on how objects move when nothing is stopping their natural path in free fall, that such an event would indeed make the cup fall (under ‘normal circumstances’). It does not matter to us that we merely assume a causal connection, that we don’t refer to the push of the electrons in the paw on the electrons of the cup and why this property exists, etc. The solution is assumed as true and satisfying until we are given a reason to doubt it. Do we have reason to doubt the HPC is a question which does not follow the general rule of why-problems having satisfying solutions?

Dealing with the category-mistake

There are some differences between a regular solution and 'the hard solution', starting with the type of explanation we can find:

We try to explain the relation between the physical and the mental as a causal relationship, and for causal relations a why question is applicable. But the physical does not cause the mental or vice versa. They are merely two sides of the same coin; the first-person perspective and the third-person perspective. An analogy: When simulating mice in a maze on a computer we can ask if changes in the ones and zeroes cause the mice to exist (not on a screen but as a computational abstraction), or if the mice cause the ones and zeroes to change. Such questions are malformed; they assume the link between the two phenomena are causal, which they are not, rather than a matter of perspective based on abstraction level.

In other words and stepping outside of the analogy: the physical and the mental are not of the same ontological stuff, so assuming there can be a causal connection between them is a category-mistake (Ryle, G. 1949).

A more appropriate question would be: What sorts of changes in ones and zeroes are interpreted as mice, and how are mice expressed in ones and zeroes. One would assume we can try to connect 'the world of consciousness' and 'the world of physical processes' like we can try to connect 'the world of mice' and 'the world of computation'. But asking why mice are ones and zeroes, or why ones and zeroes create mice, is the wrong sort of question.

The proof of impossibility

Can we then satisfy ourselves with an explanation connecting the physical and the mental? Here are the differences between the sort of problem which seeking the connection in the mice-world is, and the HPC:

1. The missing-data argument, against definitive rules.
 - The change in perspective in the mice-world scenario (and its equivalents) is caused by abstracting at a different scale.
 - The change in perspective in the HPC is caused by projection.

We imagine how we would feel given the (internal and external) circumstances of the entity in our mental model of the world. The accuracy of the projection is limited by how similar we are (or rather: can be) to that which we wish to project on. There are already inaccuracies in imagining your past self and the further from your current state somebody is the more inaccurate we can expect the projection to be. We may feel justified in saying "I know how you feel" to another human in circumstances we are familiar with, but the accuracy of such a statement plummets when we try to imagine being a different species or even for a lot of scenarios involving wildly different cultures. By this limitation we might manage to connect our own neural states to an accurate imagining of what

it is like to be in those neural states, but since we simply lack data on the experience-side for other minds we will never be able to solve the HPC to a satisfying degree; our explanation will always lack the ability to generalize, on both the intuition-side and the prediction-side. When we have a fuller explanation/connection for human consciousness and feel comfortable generalizing this to other biological creatures then we would, for example, never be able to answer whether other types of computation can also be conscious. Despite the general lack of absolute direct ontological access to the minds of others as pointed out by the skeptics, also known as 'the problem of other minds', we do not have any specific reason to doubt that a theory of consciousness doesn't hold up when applied to other humans, as there are no known fundamental differences which would restrict this freedom for generalization. But we could never be comfortable in applying this theory on a computer simulating neuronal interactions. What if consciousness is tied to how the stuff, which we describe as particles, interacts in some specific structure which we do find in brains but not computers? And this is just one of many possible doubts which can be leveled against this kind of generalization, for just one possibly functionally equivalent process.

As an example of the 'degrees in comfortableness of generalizability', which will make the concept itself more comfortably generalizable, imagine having access to the world through one window for your whole life plus some basic grasp of language. When seeing various helium balloons fly by you automatically develop the rule that balloons fly upwards. But when asked if this rule still holds if a substance which at that point you come to know as helium is no longer in the balloon you, albeit not having any way to verify your suspicions either way, begin to doubt the generalizability of your rule. You lack the ability to determine the factors responsible for the previously established rule, and therefore also lack the ability to say whether the rule holds in different situations. An almost trivial concept, but often overlooked in practice.

To illustrate the irreparability of this issue: one could imagine an attempt to bridge the gap between human and machine by removing a piece of someone's brain, scanning it into a computer, and connecting to the dendrites and synapses left behind a simulation of the removed part. The idea is then to simply ask the participant whether they are conscious of whatever that piece of brain was the neural substrate of. Intuitively this might seem to work, but we know it cannot: since the simulated part is functionally equivalent, the behavior of the participant will remain the same. No different belief will be created about their conscious states or lack thereof. In this experiment one could even only give the brain random signals, receiving none, which just happen to be identical to what the removed part would have given (including all potentially relevant changes in neurotransmitters etc.). Now there is nothing which is remotely recognizable as a potential substrate for consciousness yet the beliefs and behavior still do not

change (reductive functionalism leads inevitably to behaviorism). In either case the participant has simply become too different, too doubt-inducing, for us to still feel comfortable imagining the same sort of experience in.²

2. The computational complexity argument, against understanding.

-The changes in ones and zeroes 'beneath' mice-world are still calculable to some extent. We could, in the simplest possible case, imagine how the bits are flipped in the memory and how that information is fed to the cpu and back, etc. We can call computers physical representations of part of our reasoning. The changes in the world (in the computer) are accurately abstracted into discrete concepts and the manipulations between them are fairly linear and comprehensible.

-The same cannot be said for our brains.

Experience emerges from parallel processing and non-discrete processes. Even if we can form systems of concepts which map accurately enough to the latter, the immense number of abstractable causal links for even a tiny sliver of experience is far too great to fit into our short-term memory. This tiny part of the neural substrate for an experience needs to be simultaneously imagined due to the parallel processing involving many moving parts that are relevant simultaneously. One could object that we are just searching for one type of process common to all conscious experience and that this type does not have to be so complex. But the point is that we have to be able to logically imagine a process to be conscious and we cannot identify with a conscious process small enough to imagine. The usual procedure of abstracting complex causal processes into rules, abstracting those rules into higher rules, etc. does not work here either. It is unlike something like a mathematical proof or a system-theoretical archetype where (theoretical, abstract versions of) causal chains are made into rules, which are made into rules, etc. until only the mentioning of the name of the proof/archetype suffices for intuitive understanding why X has behavior Y. Because these products of our reasoning are linearly structured and thus consciously understandable; we can go through them step by step. The neural substrates for (a sliver of) experience however are simultaneously parallel. This would be the equivalent of understanding intuitively the large-scale behavior of an instance of Conway's Game of Life based on knowledge of the behavior of

² This serves as a counter-argument to Chalmers' notion of 'organizational invariance'. (Chalmers, 1995) He attempts to show by reductio ad absurdum that functionally equivalent systems should give rise to the same sort of experience, since swapping two functionally equivalent systems shouldn't make the experiencer notice this happens. Which is true. But that doesn't mean the experiencer has the same experience. It is quite possible, though counterintuitive, someone believes they have had a certain experience while no such thing could've happened. Think of false memories, or the feeling of having vision in the blind spot of the visual field. Lacking some experience does not imply noticing that lack.

individual pixels. Using pen and paper you might predict the next iterations, but you won't have understanding. There will be no eureka-moment for consciousness as no internal systems of concepts on their own logically compute to something recognizable as a conscious state (i.e. some element in the model of the third-person perspective of conscious processes with some recognizable functional similarity as something in the model of first-person perspective of the same conscious process).

This limits the sort of explanation we can have to one which gives general rules accurately mapping brain states to conscious states, and vice versa. This explanation will not make sense logically at the 'explanatory gap' and it cannot be used to determine with certainty the presence of conscious processes in all cases, but it will be useful for manipulation of our experience. Or to restate this conclusion: only easy problems remain as worthy endeavors.

Applications of the proof of impossibility

Searle makes a mistake addressed by the missing-data argument in his reply to 'the systems reply' in his famous Chinese room thought experiment (Searle, 1980). Searle argues that a computer cannot have a mind/understanding, so consciousness, because we can imagine someone who does not know Chinese sitting in a room and making all the computations a Turing machine would make if a computer were to simulate the brain of a Chinese speaking person while not understanding a word of Chinese. (This is the form the thought experiment has to take to account for all the complexities; in the earliest version there is simply someone looking up a correct response in a large dictionary.) 'The systems reply' states that it is not the individual who has to have understanding but it is the entire system which has understanding of the Chinese language. Searle's reply to this is that the individual in the room can, in theory, internalize all other aspects of the room to create the causal chains of the artificial brain in their brain, yet this person still lacks understanding of Chinese. The simple mistake in Searle's reasoning is that he already assumes this subsystem cannot be conscious because it is so different from the non-Chinese-understanding individual and the whole thought experiment becomes an experiment in question-begging.

But the more fundamental mistake is trying to find a way in which we can reason about the possibility of consciousness in the subsystem (or the whole room as a system). It is the missing-data argument which shows us this approach is doomed to fail: it is not possible for us to gain information on what it could possibly be like to be the system which functionally 'understands' Chinese. Searle is stopped from generalizing our human consciousness to the system because that system is so different, but a lack of knowledge does not imply the knowledge of a lack. (As the aphorism goes "absence of evidence is not evidence of absence".)

Reggia et al. (2014), who I will use as a prototypical example, are somewhat similar to me in their suggested approach: They would like to close the explanatory gap through making computer models which capture behavioral correlates of consciousness and then mapping these models onto neurocomputational states, once sufficiently complex. This they call the 'computational explanatory gap' and 'solving the easy problems' (referring to Chalmers his categories). They then suggest that when the aspects of such a model which are always present in conscious states but not in unconscious states are found, then the neural correlates for consciousness can be deduced. This I agree with, if it is assumed the addition of such an aspect needs to be present in ALL conscious states and if it is assumed this gives a functional area for which to look and not the exact 'building block' of consciousness itself since the exact role of such an aspect cannot be found using this method alone. (If these two additions are not included then one might start assuming that the reportability or language is fundamental to consciousness. These ideas obviously fail to account for all forms of experience we know we're having.) But then they suggest that "should neurocomputational correlates of consciousness be discovered, they may also provide a direct route to investigating the possibility of instantiated machine consciousness, to identifying candidate properties that could serve as objective criteria for the presence/absence of phenomenal consciousness in machines and people, and perhaps even to a better understanding of the fundamental nature of consciousness," and that the explanatory gap (and with it HPC) will fade away, equivalent to claiming that our (complete) lack of understanding what life is has dissipated by our better functional knowledge of life-processes. In both the argument and analogy they fail to recognize the consequences of the missing-data argument:

- These 'candidate properties' are only based, and can only be based, on human brains and they are always limited by the explanation a model can provide. What is to say 'property X' which is functionally equivalent to machine 'property Y' while not being similar in structurally, chemically etc. (and all other aspects we might come up with, assuming we can even fathom the necessary aspects) is generalizable in consciousness-aspect to property Y? We can only, epistemologically comfortably, generalize the consciousness-aspect if there are no potentially critical aspects dissimilar between ourselves and some other systems/processes (names which presuppose some aspects above others as well).
- The question of what life is, is based on how some concept is applied to some 'objects'(/phenomena) in our mental world, while the question of what consciousness is, is based on a concept not 'referring to'(/associated with) anything in our mental world, as in: a model of the information caused by the senses, but to the process of a mental world and all other aspects of our complete phenomenological experience itself. In the former scenario we reason about a world for which we assume there is a cause in the objective world beyond the senses, while in the latter scenario we reason about the experience itself and are thus stuck with something only directly accessible to ourselves.

Claims about when something is alive are testable for all 'things' we see, while only we ourselves (= 'the thing' which we refer to as the experiencer of the experience in our model of the world itself) as a singular 'thing,' together with all 'things' similar enough to us, to make us feel comfortable generalizing experience to (then experienced as 'person-things' (people)) are testable for when something is conscious.

As far as I have found there are no serious publications on creating understanding regarding how some processes are conscious, possibly because there are no explanations of our cognitive processes sufficiently detailed.

Non-solutions for the hard problem

Chalmers has been rightfully criticized for his reliance on intuition and metaphysical assumptions, but none of these criticisms deflate the hard problem as formulated by him.

Dennett (2018) his question "And then what happens?" doesn't bring us closer to intuitively and/or correctly projecting consciousness on some system, for reasons already discussed.

Patricia Churchland (1996) makes the same sort of analogy which Reggia et al. do regarding previously unsolved questions.

Pigliucci (2013) correctly identifies the category mistake, it does indeed make no sense to ask how experience and physical processes are causally connected, but then he posits that causal explanations are the only sort of explanations. Which is clearly false, as we can see from the mice-world analogy.

Various panpsychists and/or idealists, like Kastrup (2017), claim the hard problem is solved when the material and mental are seen as the same thing. Usually, these ideas merely reduce the ontological categories we use to predict experience on paper, while supplanting 'the physical outside world' for some supposedly and unfalsifiably mental stuff with the exact same behavior, properties, etc. as the physical stuff would have. It doesn't solve the hard problem of consciousness since it still cannot explain when and why the 'general mental stuff' gives rise to consciousness as we know it. A claimed more parsimonious 'explanation' of reality is no convincing reason to muddle our language with more words with disconnected double meanings.

Clues for the way forward

Taking this into account, we are still quite able to make predictive models which link the causal structures we can abstract from our phenomenological world to a *description* of

the world itself (which, again, must be duly noted, is a subset of the former). There are various other abstractable phenomena which limit and shape the solution space for such a predictive model even further. The following are somewhat speculative yet mostly undeniable (summaries of) observations and a resulting hypothesis showing what type of neural processes we ought to be looking for as consciousness-candidates:

- Learning complex information requires consciousness, and when something unexpected happens attention is drawn to it. This has implications for the potential function of consciousness.
- The phenomenological world we inhabit can be divided into the sensory world and the abstract world. The latter tends to inform the former more than vice versa. This world is also mostly built from experience, the sensory as well as the abstract, according to the highly influential and promising 'predictive processing' theory of the mind.
- All conscious activity is characterized by choices, even the choice to do nothing, and these decisions are guided by (un)desirable feelings. These feelings/associations depend on the relations between our phenomena, and in the abstract world these can be described as 'systems of concepts' (which formal logic systems attempt to describe).
- Tendencies, habits, reflexes, etc. occur outside of conscious processing but conscious processing is based on these; we tend to think and act in certain pre-established patterns, which are based on previous pre-established patterns, etc.

Obvious as most of these observations might be, they do reveal a structure of experience which must have a functional equivalent in the brain. By noting what aspects always appear throughout conscious experience, when their equivalents are not to be found in unconscious processes, and then noting the patterns in these phenomena themselves we can deduce the fundamental causal structures of experience. It can of course be the case that, for example, we notice phenomena A as always present while we're missing that phenomena B which is the cause of A (and actually 'the trick of consciousness') appears almost simultaneously, but it gives a clue towards finding the neural conscious processes nonetheless as their structure has to be identical to the structure of the phenomena. So the hypothesis, gotten through continually testing various possible combinations of the systems of concepts behind (among many others) the previous observations (while also combining them with basic knowledge of the mind/brain) on my own phenomenology, becomes:

Conscious processes are those which compare and choose from the desirability of the associations in our model of the world, which need to be one causally connected whole as the organism cannot practically live with constantly contradicting choices. The evolutionary 'purpose' of consciousness is to feel-choose.³ Every single experience you

³ This idea of what conscious processes are is similar to what Dennett (2018) proposes, yet independently developed. Contrary to Dennett however, I do not believe this is a beginning of the dissolution of the HPC, as should be clear from the rest of the paper. The 'theory of

have is characterized by this. There would be no 'use' to experience if you didn't have some judgement and decision on it. The (un)desirability of some association is the same as our level of aversion/eagerness to flow into the state predicted by the association. (A following temporal state, not necessarily spatial in our web of associations.) The concept of 'an active choice' implies energy expenditure, but when that part is ignored we must conclude we continually choose with just varying degrees of deliberation, focus and energy. The deliberation, and we of course also choose to deliberate, is done in the abstract world and involves navigating through our model of the world. This model of the world is a network of concepts and their relations. Some of these are held to be true, conditional on other concepts and relations, and/or felt with some other association. Based on what is felt to be true (or appropriate, beautiful, warm, satisfying, etc.) deliberation is stopped and a plan is chosen. You may notice that the abstracted versions of these processes ('plan', 'stopping', 'deliberation', etc.) are already a step too far removed from the actual experience and thus always fail to properly encapsulate it. The lived experience is continuous and the associations come in degrees.

Strong prediction-errors, as defined by the predictive processing framework, and strong emotional associations (among many others) automatically call our attention because that is where consciousness is needed to make some decision (phenomenologically, the desirability-aspect is an alleviation of a sense of wrongness); the fast processing gives control to the slow but adaptable processing. The point at which the parallel system 1 processes, well explained by the predictive processing framework, combine into serial system 2 processing (terminology from Kahneman, 2011) is the point before which these conscious processes happen. And we learn through the activity left behind in these processes, in the short and long-term.

To accurately correlate and thereby prove, in so far as possible, what these processes are neurologically, is the following step in this reasoning process which at the very least serves as a template for further questioning. But I encourage the reader to build on this basic framework if they feel so inclined, as the consistent patterns which can be abstracted can also, for example, be used to discover the philosophical consequences: Our feelings about the processes in the(/our internal) world include epistemological and ethical processes, if we see ethics as the question of generalizing principles of desirability. The perspectivist and internalist definition of ethics is a result of the necessity to make philosophy coherent with the well-predicting scientific theories of the world, introduced in the beginning of the paper. Notice some epistemological process happening here which we can call "epistemological hierarchy".

Finding the patterns responsible for these epistemological processes is what is both part of the aim of the new search for the fundamental building blocks of consciousness and the perspectivistly coherent way of doing philosophy. (The same can be said for other areas of philosophy which also, actually, deal with how the mind functions.) This

consciousness' presented here is also further developed, but otherwise completely in line with Dennett's idea. So for more arguments supporting the idea I refer to his paper.

quest will lead to an amalgamation of all the currently popular attempts at describing the principle behind 'true truth', which have been likened to the parable of blind men each describing one part of an elephant, when it is more concerned with making an abstract model mapping to and predicting when we have 'truth-feelings' rather than getting to a simple explanation.

Conclusion

I have shown, from a phenomenologist, perspectivist and internalist philosophical framework, that the HPC as it is commonly understood firstly needs to be reformulated to even make sense in our current best working materialist view of the world and even then the reformulated version will never be answered completely or intuitively. After cutting off this branch of philosophy of mind I then sprouted a new, sturdier one by concluding only the 'easy problems' remain and by composing a foundation for answering them.

References

- Chalmers, D. (1995). "Facing up to the problem of consciousness". *Journal of Consciousness Studies*, 2: Page 200–19.
- Churchland, P.S. (1996). The hornswoggle problem. *Journal of Consciousness Studies* 3 (5-6): 402-8.
- Dennett D. C. (2018). Facing up to the hard question of consciousness. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 373(1755), 20170342. <https://doi.org/10.1098/rstb.2017.0342>
- Kahneman, D. (2011). *Thinking, fast and slow*. New York : Farrar, Straus and Giroux
- Kastrup, B. (2017). The Quest to Solve Problems That Don't Exist: Thought Artifacts in Contemporary Ontology. *Studia Humana* 6 (4):45-51.
- Pigliucci, M. (2013). What hard problem? *Philosophy Now* (99).
- Reggia, J., Monner, D. & Sylvester, J. (2014). The Computational Explanatory Gap. *Journal of Consciousness Studies*. 21.
- Ryle, G. (1949). "The concept of mind". London ; New York : Hutchinson's University Library. Page 16.
- Searle, J. (1980). 'Minds, Brains, and Programs', *Behavioral and Brain Sciences*. 3. Page 240