

A Key Hidden in Plain View

Nir Aides — nir@winpdb.org

(An unpublished paper written in July 2016)

Philosophers have been contemplating the nature of the mind for centuries and have produced mountains of intricate jargon, thought experiments, and views, that map a landscape of interminable disputes.¹ One such dispute is between philosophers who believe that the mind can be explained as a mechanism and philosophers who insist it cannot. In this paper I take a look at this dispute and argue that it is unique in philosophy and a key to the nature of the mind.

Q(core): Can the mind be explained as a mechanism?

And one ought to wonder:

Q(meta): Why do some people believe that the mind can be explained as a mechanism and others insist it cannot?

One might argue like René Descartes (1637) that there is no point in philosophy which is not disputed — and that there are many reasons why philosophers might answer Q(core) one way or the other.

The mind is a motherlode of problems and concepts over which people can argue indefinitely — free will, genuine creativity, love and emotions, the self and the problem of identity and more.

And we use some of these concepts without agreement on a definition, because they are incoherent or because, as Wittgenstein argues in *Philosophical Investigations* (PI, §77), their boundaries are blurred.

Imagine a poet who says that the mind cannot be a mechanism because a mechanism cannot love, or a philosopher who argues that the mind cannot be a mechanism because a mechanism cannot have libertarian free will.

¹ The abundance of views and intricate terminology in philosophy of mind reminds me of a scene in Monty Python's *Life of Brian*, in which Brian meets the People's Front of Judea for the first time: "Brian: Excuse me. Are you the Judean People's Front? Reg: Fuck off! We're the People's Front of Judea" — <https://www.youtube.com/watch?v=a0BpfwazhUA>

Philosophers are generally not in agreement even on central concepts in philosophy of mind, such as physicalism (Chomsky, 2009), consciousness, and qualia.

People may be biased toward a religious belief or a scientific worldview — like a religious person who dogmatically insists that a mechanism cannot have a soul, or a scientist who insists on ignoring anything that cannot be measured.

There are many such reasons but I believe that it is just the circumstance that enabled a fascinating phenomenon to remain hidden in plain view for so long.

Wittgenstein says “Look at the blue of the sky and say to yourself, ‘How blue the sky is!’ — When you do it spontaneously — without philosophical purposes — the idea never crosses your mind that this impression of colour belongs only to you. And you have no qualms about exclaiming thus to another. And if you point at anything as you say the words, it is at the sky. I mean: you don’t have the pointing-into-yourself feeling that often accompanies ‘naming sensations’ when one is thinking about the ‘private language’. Nor do you think that really you ought to point at the colour not with your hand, but with your attention.” (PI, §275)

Right, but when I look at the blue of the sky, I am often aware of its private aspect, and I point at it, so to speak, with my attention, and I note it is utterly mysterious and cannot possibly be the product of a mechanism.

Q(blue): As you stare at a blue thing, do you see anything in your mind that may not be explained in mechanical terms?

The thing I am speaking of is sometimes called qualia, which philosophers roughly define as “what it is like” to see the blue of the sky or “the way things seem to us”. But these definitions are problematic because they may be taken to be said about cognitive processes that may be conceivably explained mechanically, such as associations, judgements, memory, and so on.

And this is a problem that will come back again and again to bite us — that people may be using the *same words* to talk about *different things*.

One might say that seeing the blue of the sky is like jumping into a beautiful abyss, or that the blue of the sky seems different today.

But the thing I am speaking of does not seem like anything at all, and there is nothing it is like.

Wittgenstein says “It’s not a Something, but not a Nothing either! The conclusion was only that a Nothing would render the same service as a Something about which nothing could be said.” (PI, §304)²

Wittgenstein argues that we have no criteria of identity or correctness for it (PI, §253 - §279) and that it transcends the psychological self (§398). However, much of Wittgenstein’s criticism applies to introspection of cognitive processes as well, and scientists may argue that one day they may have better access to such processes than introspection affords.

One would like to say that qualia are about the quality rather than quantity of a perception — that they designate that which cannot be described.

But, we can imagine someone who says that the color of the sky has a blue quality to it, because he is not in a position to quantify the color through introspection, though he could have been in principle — like a violin player with absolute pitch who can identify a note, or an omnipotent wine expert who can consistently describe the bouquet of a wine with detailed, fantastic associations, while you can only report that it seems indescribably nice.

And that is the curse of a Something about which nothing can be said, for whatever you say about it may be taken to be said about something else.

And yet some people insist that even in simply looking at the blue of the sky, they find a mysterious and non-mechanical not-a-Nothing.

² This reminds me of how Osho describes meditation: “It is very difficult to verbalize it. To say something about meditation is a contradiction in terms. It is something which you can have, which you can be, but by its very nature you cannot say what it is.” Incidentally, he argues that people have come to use the word meditation to talk about something entirely different, akin to an activity involving concentration or contemplation.

Daniel Dennett (1993, pp. 82-85), who believes that the mind may be explained as a mechanism, compares such people to tribesmen who believe in a blue-eyed god of the forest. In an interview at the Moscow Center for Consciousness Studies (2012), Vadim Vasilyev asks him “How do you conceive this extra consciousness which actually doesn’t exist, but which can be conceived?” and Dennett answers “I think that a lot of people, they might not admit it but I think, when they think of consciousness in this extra sense, they imagine a sort of radioactive glow.”

Someone could say “I can see in my mind the thing of which you speak, but I owe you no explanation for a Something about which nothing can be said.”

But Dennett (1993, p. 372) flatly denies the existence of the thing I am speaking of, and my view, put bluntly, is that we should take him at his word — but other philosophers disagree.

When Dennett (1979) writes “My view, put bluntly, is that there is no phenomenological manifold in any such relation to our reports. There are the public reports we issue, and then there are the episodes of our propositional awareness, our judgments, and then there is — so far as introspection is concerned — darkness.”³ David Chalmers (1997a, p. 190) resolves it as a linguistic dispute: “What might be going on when someone claims that introspection reveals only judgments? Perhaps Dennett is a zombie. Perhaps he means something unusual by ‘judgment.’ Most likely, however, he has taken something else for introspection: what we might call extrospection, the process of observing one’s own cognitive mechanisms ‘from the outside,’ as it were, and reflecting on what is going on.”

But if we consider the entire edifice of arguments philosophers erect to support their views, we are likely to find arguments we dispute or concepts we use differently. Then we may argue that we disagree because of these misunderstandings which we are not likely to resolve.

³ In case you are left wondering what Dennett means by “any such relation”, then it has something to do with “episodes [that] are the momentary, wordless thinkings or convictions (sometimes misleadingly called conscious or episodic beliefs) that are often supposed to be the executive bridges leading to our public, worded introspective reports from our perusal or enjoyment of the phenomenological manifold our reports are about.” — good luck making sense out of that.

My intention in focusing the question on an introspective staring at a blue thing is to isolate the discussion as much as possible from these theoretical edifices, blurred concepts, and cognitive capacities, within the circumference of the problem.

Naturally, as we stare at a blue thing our mental activity does not stop, and we might wander off on a stream of thought and end up manifesting all of these capacities. Nevertheless, we may say that Q(blue) is concerned with an empty-minded visual meditation, even if it only manifests in the form of an occasional moment of grace.

Presumably, during such a moment we would still have an experience of color — at least some people insist that such is the case.

Dennett (2003) and other philosophers and scientists argue that introspection is a poor source of evidence and that we can be embarrassingly mistaken about our inner world — consider for example change blindness, which occurs when an observer fails to notice significant changes introduced into a visual scene.

If someone metaphorically pointed his finger at qualia and reported “look, there it is”, we would presumably have to inform him that his finger is a bag of tricks that cannot be used properly as a pointing device. The unlucky subject would then protest in vain that qualia would still be there even if it turned out that his finger was really a banana.⁴

Despite the curtains of cognitive illusions, it is in introspection where one might find qualia if they exists at all to be found, not in philosophy books or columns of numbers.

In an interview at the Moscow Center for Consciousness Studies (2013), Dmitry Volkov asks Chalmers why most philosophers are physicalists, and Chalmers answers that “the progress of science has led many to think that the world must be physicalistic” and that “although many people think that physicalism has problems, the alternatives also have serious problems.”

⁴ Bruce Lee explains it to a young kung fu student in the movie Enter the Dragon: “It’s like a finger pointing away to the moon. Don’t concentrate on the finger or you will miss all that heavenly glory.” — <https://www.youtube.com/watch?v=CwPWDMvT21E>

John Searle (1997, p. xiii) writes that philosophers have an urge for reductionism and materialism because they suppose that otherwise they will somehow reject the scientific worldview.

Dennett (1993, p. 37) appears to confirm this explanation when he writes: "This fundamentally antiscientific stance of dualism is, to my mind, its most disqualifying feature and is the reason why in this book I adopt the apparently dogmatic rule that dualism is to be avoided at all costs."

But one may still wonder why Dennett can be so dogmatic about the scientific worldview in the first place.

Consider this conversation between Einstein and Rudolf Carnap, as reported by Carnap (1963): "Once Einstein said that the problem of the Now worried him seriously. He explained that the experience of the Now means something special for man, something essentially different from the past and the future, but that this important difference does not and cannot occur within physics. That this experience cannot be grasped by science seemed to him a matter of painful but inevitable resignation. I remarked that all that occurs objectively can be described in science; on the one hand the temporal sequence of events is described in physics; and, on the other hand, the peculiarities of man's experiences with respect to time, including his different attitude towards past, present, and future, can be described and (in principle) explained in psychology. But Einstein thought that these scientific descriptions cannot possibly satisfy our human needs; that there is something essential about the Now which is just outside the realm of science. We both agreed that this was not a question of a defect for which science could be blamed, as Bergson thought. I did not wish to press the point, because I wanted primarily to understand his personal attitude to the problem rather than to clarify the theoretical situation. But I definitely had the impression that Einstein's thinking on this point involved a lack of distinction between experience and knowledge. Since science in principle can say all that can be said, there is no unanswerable question left. But though there is no theoretical question left, there is still the common human emotional experience, which is sometimes disturbing for special psychological reasons."

It is not clear if Carnap is struggling to understand Einstein, or simply believes that one need not explain a Something about which nothing can be

said. But it seems reasonable to interpret Einstein as speaking of the problem at hand, and if that is the case, then one may wonder why Dennett insists on being more Catholic than the Pope about science.

Could Dennett's scientific dogmatism be a symptom rather than a cause?

We can make Q(core) more recognizable by expressing it in terms of computation rather than mechanisms, for we can easily imagine the body of an intelligent robot driven by a system of cogwheels, but not so its brain — for that, we might think, it needs a computer.

One way to think about the nature of the mind is the Computational Theory of Mind, according to which, the mind is a form of computation, as we understand computation since Alan Turing — as a mechanical process (Horst, 2015). This theory enables us to finally express the core disagreement in terms of a simple and familiar problem.

Imagine that scientists have built a computing system that perfectly emulates a human brain. Imagine that we take it outside and ask it to stare at the blue sky, and ask yourself the following question:

Q(ai): As you stare at a blue thing, do you see anything in your mind that you would expect to be missing in principle from a computer that stares at a blue thing?

Philosophers and scientists have answered this question in the past. Dennett (1993, p. 281) declares in his book *Consciousness Explained* that computation may be conscious in the fullest sense, and Marvin Minsky, who won the Turing award in 1969 and co-founded the MIT artificial intelligence laboratory, makes the same claim in a *Closer to Truth* interview with Robert Lawrence Kuhn (2011). On the other side of the trench you can find Searle, armed with a shotgun⁵ and the firm belief that computation cannot amount to a mind, and the physicist Roger Penrose who believes that the mind is incomputable.

⁵ Searle defending his views with a shotgun in his hands:
<http://www.newphilosopher.com/articles/john-searle-it-upsets-me-when-i-read-the-nonsense-written-by-my-contemporaries/>

But the sight of philosophers eternally arguing about such questions is as ordinary as the sight of an apple falling to the ground, and so it was largely ignored. No one in philosophy of mind is seriously bothered by Q(meta).

Computation is a marvelous tree — its roots are planted in the mechanical and its leaves reach the doors of perception.

René Descartes, the mathematician, scientist, and philosopher of the 17th century, often credited as the father of modern philosophy, believed that all material bodies operate on mechanical principles (Hatfield, 2014). But despite his wish to explain the world as a mechanism, Descartes, arguably the philosopher most identified with dualism, believed that a mechanism is insufficient for a mind, and in 1637 he argues in the *Discourse on Method* that mechanical machines “could never use words, or put together signs, as we do in order to declare our thoughts to others.” (Oppy, 2011).

Since Newton we have abandoned the hope of attaining an intelligible mechanical explanation of the world (Chomsky, 2009). Our current conception of the physical is what Leibniz would call “occult and unintelligible” (Janiak, 2014).⁶ It includes fields that permeate empty space, relativity of simultaneity, entanglement, the wave-particle duality, and the measurement problem, just to begin with — and if that is true of sunlight or an apple falling to the ground, what hope is there for the mind?

Three hundred years after the *Discourse*, Alan Turing (1936) invents the Universal Turing Machine, a simple and abstract mechanical machine that according to the Church-Turing thesis can carry out any effective computation. Then, in *Computing Machinery and Intelligence*, Turing (1950) considers the question “Can machines think?” and proposes to replace it with an imitation game, which came to be known as the Turing Test, and in which an interrogator is required to determine if a hidden interlocutor is man or machine. Turing essentially argues that Descartes was wrong, and that computers may one day respond appropriately and meaningfully to whatever is said in their presence — that a machine reducible to a mechanism could pass the imitation game for a thinking thing. Would that thing amount to a mind?

⁶ The physicist Richard Feynman famously said: “I think I can safely say that nobody understands quantum mechanics.” “Do not keep saying to yourself, if you can possibly avoid it, ‘But how can it be like that?... Nobody knows how it can be like that.’”

Turing does not give a decisive answer to that question. He writes: “In considering the functions of the mind or the brain we find certain operations which we can explain in purely mechanical terms. This we say does not correspond to the real mind: it is a sort of skin which we must strip off if we are to find the real mind. But then in what remains we find a further skin to be stripped off, and so on. Proceeding in this way do we ever come to the ‘real’ mind, or do we eventually come to the skin which has nothing in it? In the latter case the whole mind is mechanical.”

Computers cannot yet pass the imitation game, but researchers are making continuous progress in artificial intelligence. Sixty-six years after Turing’s paper, you can ask Google such questions as “When was the Discourse on Method written?” by voice, and get an appropriate and meaningful answer; computers can use words to describe images as we do (Karpathy, 2015)⁷ and learn to play video games, uncannily well (Wired, 2015).

We are gradually coming to accept it as natural whenever computation is applied successfully to problems that belong naturally in the domain of human cognition. The belief that a computer may never pass the Turing test starts to look like a case of a shrinking God-of-the-gaps.

It is possible that Descartes underestimated the capacity of mechanisms, but was that key to his belief in dualism? Had he known then what we know today of computers, would he have repented and joined the Order of Computational Theory of Mind, or would he have remained adamant?

A generation after Descartes, Leibniz (1714; 1898) writes the beautiful mill argument: “Suppose that there be a machine, the structure of which produces thinking, feeling, and perceiving; imagine this machine enlarged but preserving the same proportions, so that you could enter it as if it were a mill. This being supposed, you might visit its inside; but what would you observe there? Nothing but parts which push and move each other, and never anything that could explain perception.” Is Leibniz simply underestimating the capacity of mechanisms too?

⁷ <http://cs.stanford.edu/people/karpathy/deepimagesent/>

Leibniz's mill argument captures the essence of that which David Chalmers (1995) calls the hard problem of consciousness — that there appears to be a Something in our minds that is inexplicable on mechanical grounds. Many proponents of the Computational Theory of Mind deny the hard problem of consciousness.

A common argument by people who appeal to a physical version of the Church-Turing thesis, is roughly that the universe is computable and therefore a computer may in principle simulate a brain as a physical system. On what grounds, they ask, may we deny that such a simulation harbors a mind?

The funny xkcd comic strip *A Bunch of Rocks*,⁸ takes this idea to its extreme depicting our universe as a computation that a Sisyphus-like protagonist carries out by systematically moving rocks on an infinite stretch of sand on which he is stuck for eternity. Since one can emulate a Turing machine by moving rocks, it is possible in principle to carry out any effective computation this way.

Philosophers sometimes criticize such arguments as *loys of funny instantiation* (Maudlin) or *intuition pumps* (Dennett), complaining that they unfairly engage our intuitions, but I am actually interested in that intuitive response.

Surprisingly, plenty of people with whom I have spoken calmly insist that a bunch of rocks may be just as conscious as they are, if the rocks are moved around properly.

Some argue that while the mind may be explained as computation, it is nevertheless a computation that may not be adequately carried out by a Turing Machine — because the human brain is massively parallel, or analog, or interactive, and so on, while a Turing machine is serial, and discrete, and is meant to compute an output and halt. Therefore, to simulate what the brain does as a computing system, we would supposedly need a special architecture — for example, a massively parallel architecture, or some other yet-unheard-of architecture.

⁸ <https://xkcd.com/505/>

Indeed, the brain does not resemble a Turing machine, and one may claim that another computation model may represent better what the brain does as a computing system, but ultimately a Turing machine can simulate any computing system to arbitrary accuracy.

Others argue that a simulation of a thing is not the real thing. Searle (1984, p. 37) writes: “no one supposes that a computer simulation of a storm will leave us all wet”, to which Eliezer Yudkowsky (2010) responds by asking: “Can you have simulated information that is not really information? Can you have correct answers which are only simulated correct answers?”

There are ongoing disputes about what counts as an implementation of a computation, whether computation is observer-relative or not, and about the role played by causation in a computation that is being carried out — but we can sidestep these problems, since I do not mind granting computational functionalists a good part of what they want — an infinity of computations that are practically indistinguishable from brains in passing for a human.

There is one more source of confusion we need to consider — naive realism, which is roughly the belief that the blue of the sky is a property of the sky. People may believe that because it is a commonsense naive belief, or because they have taken this position after philosophical consideration.

It is no wonder that if someone believes that the blue of the sky is not a thing to be explained in one’s mind, the problem would “go away”. But naive realism is generally rejected by philosophers (BonJour, 2007).

It seems rather that the world we experience around us — the world we inhabit, with its colors, smells, and sounds, is the product of our brains. It is caused by external reality but it is not external reality itself, not directly. The neuroscientist Rodolfo Llinás expressed it colorfully by saying that “Life is nothing but a dream guided by the senses” (Revonsuo, 1995), and this is what Dennett (1993, p. 372) means when he writes “it really does seem as if science has shown us that the colors can’t be out there, and hence must be in here.”

One powerful illustration of this idea is Benham’s Top — a spinning disk with black and white stripes that creates the sensation of color. Where is the color? is it in the mind or in the external world?

But the point is not to convince anyone or to win an argument — it is rather to find people whose answers to Q(ai) create a crisis — people who agree on most everything except for a single intuition which drives their disagreement. Let's call such people 'peers on Q(ai)'.

If someone insists on naive realism, or denies that computation is a mechanical process, or believes that computers may never show initiative "just because", then I would not know what to make of his answer to Q(ai).

We can see that simple as Q(blue) and Q(ai) seem to be, they nevertheless require some familiarity with philosophy and computer science, and therefore, contrary to standard practice in experimental philosophy, it may not be useful to present them to random people.

However, once all is said and done the questions are simple, and although I employ "intuition pumps" to shake people off computational functionalism, the answer I almost always get is a remarkable and resolute No — the mind is a mechanism, there would be nothing missing in principle from a computer running the right computation.

I believe that human-level general artificial intelligence may be possible and that computers may one day reliably pass the Turing test, but nevertheless, I find something in my mind that cannot be explained away as computation. In my mind that thing is as evident as a sun burning in the middle of a clear blue sky, and therefore I find the general response quite amazing.

I have informally talked in person with about twenty technically informed people who appeared to be my peers on Q(ai). Almost all of them insisted that they see nothing in their mind which would seem to lie beyond the reach of computation.⁹ Some of these people are friends with whom I have discussed this subject many times over the years, and yet, I was not able to make them see that sun burning in the middle of the sky.

⁹ There were three exceptions. One person was sympathetic to Chalmers's nonreductive functionalism, another was undecided, and one reported seeing the problem after several years of conversations in parallel with becoming religious.

This general response is supposedly contrary to conventional wisdom, according to which most people find functional accounts of the mind counterintuitive (Chalmers, 1997b; Dennett, 1993, p. 226).

But Q(core) is buried under so much confusion that it renders conventional wisdom practically meaningless. There are plenty of “wrong” reasons to reject materialism and the Computational Theory of Mind.

In fact, it appears to me that most philosophically informed people embrace the idea that the mind is a form of computation.¹⁰ Consider the following opinion of famous scholars and scientists, some of whom have shown lifelong interest in the problem of consciousness.

Richard Dawkins says in an interview (Humphrey, 1986): “Could you ever build a machine that was conscious? I know that I'm conscious, I know that I'm a machine, therefore it seems to me — and I know there's nothing special — I mean, perhaps I could say I have faith, but I think, I almost know that there's nothing in my brain that couldn't in principle be simulated in a computer, so if you took the extreme policy of building a computer that was an exact simulation of a human brain, doing everything that a human brain does, point for point, mapping from human brain anatomy to computer hardware, then of course such a machine would have to be conscious in just the same sense as I know that I'm conscious.”

In a recent dialog the physicist Lawrence Krauss asks Noam Chomsky (2015) if he believes that the mind is in principle computable and could be emulated on other substrates like silicon. Chomsky, who is known as a mystician, nevertheless answers that the mind is organized matter and that “we don't know of any physical reason to believe that the particular components of that organized matter are critical for its operation”, “therefore it could be emulated presumably in some other substances.” Krauss agrees and says “We are a computing machine of some sort.” “It's hard to imagine that if we were able to reproduce all the information we wouldn't get the same person.” Chomsky then goes on to dismiss Chalmers's Hard Problem.

¹⁰ Your best chance to find a dualist is in a philosophy department.

Google's director of engineering and futurist Ray Kurzweil (2012) writes: "a computer that is successfully emulating the complexity of a human brain would also have the same emergent consciousness as a human."

The physicist David Deutsch (2004), known for the Church-Turing-Deutsch principle and for his contribution to quantum computing, writes: "there is every reason to believe that the brain is a universal classical computer."

The physicist Stephen Hawking (2010) writes: "I think the brain is essentially a computer and consciousness is like a computer program." "Theoretically, it could be re-created on a neural network".

In contrast, I have a forceful intuition that the mysterious Something that I find in my mind cannot be produced by a mechanism. I call it an intuition because it is a "clear seeming" that cannot be justified or reasoned, and it is forceful because I can no more dismiss it than any of Euclid's axioms.

I believe that it is the intuition that drives some people to Chalmers's formulation of the Hard Problem of Consciousness, and I shall therefore call it the Hard Intuition of Consciousness.

I have come to believe that most people do not have that intuition — they seem to be blind to a most evident thing — and when they do point at an intuition, they often seem to point at the wrong kind.

Stanislas Dehaene (2014), who believes that the mind may be explained as computation argues that "the hard problem just seems hard because it engages ill-defined intuitions. Once our intuition is educated by cognitive neuroscience and computer simulations, Chalmers's hard problem will evaporate."

We have many intuitions, and while philosophers employ intuitions in their reasoning, they know that intuitions can be illusory, in that they may generate false beliefs.

There are studies in experimental philosophy (Sytsma 2010) that suggest people are more inclined to intuitively attribute seeing color to a robot than feeling pain. We can easily imagine that such intuitive judgements may be influenced by naive realism, by the way the words *seeing* and *feeling pain*

are used in everyday language, and by the way robots have been depicted in science fiction over the years, and we can imagine these intuitive judgements may change in the future.

The sliding block puzzle Huarong Dao may initially invoke the intuition that it is impossible to solve. Similarly, people may intuitively believe that the mind is too complex to be simulated in principle by a computer, and we can imagine that familiarity with computation may dissolve that intuition.

I have an intuition that I have free will; however, I am quite aware that I cannot intelligibly reconcile it with determinism or indeterminism, and after philosophical consideration I concede that it may be an illusion.

Why? Because I am not conscious of the origin of my thoughts and decisions. It is as if I have an intuition about a thing that happens behind a closed door — an intuition about something that I cannot observe.

But the Hard Intuition is different — it is about a mysterious, indescribable thing which seems to be in plain view, and the closer I look, the more veridical that intuition seems to be.

In a way, it is like being able to see God and knowing that one cannot explain it as a combination of cogwheels.

It is similar in force to a geometric intuition and no amount of contemplation will dissolve it. It is up there with my intuitions of existence and time.

My intuition that something exists generates my most certain belief. Philosophers may complain that Descartes's *cogito* is logically suspect or defective but it would not make the slightest difference. It is a finger that points at truth.

My intuition of time is forceful as well. There seems to be a Now and there seems to be continuous change. I am aware that this intuition is unintelligible, but I cannot explain it away.

If someone were to insist that nothing really exists or that nothing ever changes, I would have to tell him "good luck with that."

Someone could argue that the Hard Intuition is the result of a catastrophic defect in my thinking — I know from experience that in my dreams I may fall in love with a gorgeous woman and never notice that she literally has the face of a goldfish, or carefully study a triangle and conclude that it has four edges;¹¹ and I know from experience that even in my waking life I can make incredible errors of judgment. But the Hard Intuition is a most persistent seeming that never falters, and I suspect that the force of the *cogito* and my intuition of time flows from it.

Dennett (2001) writes: “There is a powerful and ubiquitous intuition that computational, mechanistic models of consciousness, of the sort we naturalists favor, must leave something out — something important. Just what must they leave out? The critics have found that it’s hard to say, exactly: qualia, feelings, emotions, the what-it’s-likeness (Nagel) or the ontological subjectivity (Searle) of consciousness.”

Does Dennett see the problem after all or does he merely present the position of his colleagues?

Dennett points out their difficulty in saying what is left out — a point that has been the source of endless debates and confusion, for it is impossible to say what is left out. In addition, I believe that the Hard Intuition is not ubiquitous at all.

He goes on to define philosophical zombies as beings that are “behaviorally indistinguishable from a normal human being but utterly lacking in consciousness”, and he expresses the intuition as “the conviction that there is a real difference between a conscious person and a perfect zombie — let’s call that intuition the Zombic Hunch — leading [some philosophers] to the thesis of Zombism: that the fundamental flaw in any mechanistic theory of consciousness is that it cannot account for this important difference.” He

¹¹ Richard Feynman comically recounts how he came to abandon his practice of Lucid Dreaming for similar reasons: “I’m dreaming one night as usual, making observations, and I see on the wall in front of me a pennant. I answer for the twenty-fifth time, ‘Yes, I’m dreaming in color,’ and then I realize that I’ve been sleeping with the back of my head against a brass rod. I put my hand behind my head and I feel that the back of my head is soft. I think, ‘Aha! That’s why I’ve been able to make all these observations in my dreams: the brass rod has disturbed my visual cortex. All I have to do is sleep with a brass rod under my head, and I can make these observations any time I want. So I think I’ll stop making observations on this one, and go into deeper sleep.’ When I woke up later, there was no brass rod, nor was the back of my head soft.”

concludes: “A hundred years from now, I expect this claim will be scarcely credible, but let the record show that in 1999, John Searle, David Chalmers, Colin McGinn, Joseph Levine and many other philosophers of mind don’t just feel the tug of the Zombic Hunch (I can feel the tug as well as anybody), they credit it.”

Dennett’s view becomes clear if we imagine his zombies as future intelligent computing systems. Other philosophers define philosophical zombies as physically identical to humans (Kirk, 2015) but in that form zombies become irrelevant unless one presupposes that reality in general and brains in particular are computable, and that is an open question.

David Deutsch (2004) writes: “The most straightforward such idea, and also the most extreme, is that the whole of what we usually think of as reality is merely a program running on a gigantic computer – a Great Simulator.” but he rules it out since “It is in the very nature of computational universality that if we and our world were composed of software, we should have no means of understanding the real physics – the physics underlying the hardware of the Great Simulator itself.”

Ultimate reality cannot be a computation since computation, as we understand the concept, requires a computer and that entails an infinite regress.

But what if we restrict ourselves to observable phenomena? The physicist Seth Lloyd (2013) writes that “all observed phenomena are consistent with the model in which the universe is a quantum computer”. But even if that is true, and leaving aside what some people, including Einstein, claim about the human mind, we cannot rule out the possibility that such phenomena will be discovered or are already in plain view.

In *The Character of Physical Law*, the physicist Richard Feynman (1967, p. 33) says: “[The law of gravity] is not exact; Einstein had to modify it, and we know it is not quite right yet, because we have still to put the quantum theory in. That is the same with all our other laws — they are not exact. There is always an edge of mystery, always a place where we have some fiddling around to do yet. This may or may not be a property of Nature, but it certainly is common to all the laws as we know them today.”

Later, in *Simulating Physics with Computers*, Feynman (1982) notes: “A very interesting question is the origin of the probabilities in quantum mechanics.”

Consider a pair of entangled photons. According to Einstein (1935) one of the particles may be used to predict with certainty the value of a physical quantity of the other particle. But according to physics, the only way to compute such a prediction would be to use the first particle as a black box or a Turing Oracle.

Scientists and philosophers have proposed various explanations for this phenomenon including superdeterminism, superluminal particles, many-worlds, and more recently wormholes that connect entangled particles (Jensen, 2013). But another interpretation remains — that it is an observable incomputable phenomenon.

The Hard Intuition compels me to believe that the mysterious not-a-Nothing in my mind is not mechanical.

Evolution compels me to believe that it is not epiphenomenal — it is causal even if I cannot conceive how.

That, in turn, compels me to believe that my ability to observe and report it, is a case of causality rather than a convoluted case of epiphenomenalism.

And that compels me to believe that there is no mechanical explanation of my behavior and that brains are not reducible to computing systems.

Back to Dennett (2001) — while discussing Leibniz’s mill, he appears to frame his own Zombic Hunch in terms of complexity by asking “Might it be that somehow the organization of all the parts which work one upon another yields consciousness as an emergent product?” “We have learned how to think fluently and reliably about the cumulative effects of intricate cascades of micromechanisms, trillions upon trillions of events of billions of types, interacting on dozens of levels. Can we harness these new powers of disciplined imagination to the task of climbing out of Leibniz’s mill?” And he concludes: “We are quite certain that a naturalistic, mechanistic explanation of consciousness is not just possible; it is fast becoming actual.”

The last man on earth sat alone in a room. There was a knock on the door...¹² First, it filled his heart with hope, but then he remembered — every evening as the temperatures dropped, a crooked spring in the door's internal mechanism snapped and triggered a cascade of clicks.

Then he put a record on the gramophone, and as the sound of violin turned his room into heaven, he tried to eat the music with a spoon.

Metaphorically, philosophers and scientists like Dennett and Dehaene try to explain the mechanism of locked doors, but to some people, they make as much sense as the lonely man dancing around with his spoon.

“The aspects of things that are most important for us are hidden because of their simplicity and familiarity. (One is unable to notice something — because it is always before one's eyes.) The real foundations of their inquiry do not strike people at all. Unless that fact has at some time struck them. — And this means: we fail to be struck by what, once seen, is most striking and most powerful.” (Wittgenstein, PI, §129)

The 2500-year-old Hindu Upanishads tell of a mysterious internal reality that most people do not see — it is called ‘Atman’, usually translated into English as the *Self* with a capital S to distinguish it from the psychological self. The Upanishads identify Atman with another concept called ‘Brahman’ that stands for the nature of ultimate reality.

“The Self is the ear of the ear, mind of the mind, speech of speech. He is also the breath of the breath, and eye of the eye.” (Kena Upanishad)

“To many it is not given to hear of the Self. Many, though they hear of it, do not understand it.” (Katha Upanishad)

“Subtler than the subtlest is this Self, and beyond all logic.” (Katha Upanishad)

“The Self is not known through study of the scriptures, nor through subtlety of the intellect, nor through much learning. But by him who longs for him is he known.” (Katha Upanishad)

¹² These two famous lines first appeared in Fredric Brown's short story ‘Knock’, published in 1949.

“This Brahman, this Self, deep-hidden in all beings, is not revealed to all; but to the seers, pure in heart, concentrated in mind — to them is he revealed.” (Katha Upanishad)

“He truly knows Brahman who knows him as beyond knowledge; he who thinks that he knows, knows not. The ignorant think that Brahman is known, but the wise know him to be beyond knowledge.” (Kena Upanishad)

“Soundless, formless, intangible, undying, tasteless, odorless, without beginning, without end, eternal, immutable, beyond nature, is the Self.” (Katha Upanishad)

“The Self-Existent made the senses turn outward. Accordingly, man looks toward what is without, and sees not what is within. Rare is he who, longing for immortality, shuts his eyes to what is without and beholds the Self.” (Katha Upanishad)

“He through whom man sees, tastes, smells, hears, feels, and enjoys, is the omniscient Lord.” (Katha Upanishad)

“In one’s own soul Brahman is realized clearly, as if seen in a mirror.” (Katha Upanishad)

Ironically, Dennett is not far off the mark when he compares dualists to tribesmen who believe in a blue-eyed god of the forest. The Bhagavad Gita narrates the dialogue between Prince Arjuna and his charioteer, who reveals himself as Lord Krishna, the blue-colored Hindu Deity representing Atman, the Self, and Brahman, the ultimate nature of reality.

“The glory of the Self is beheld by a few, and a few describe it; a few listen, but many without understanding.” (§1.29)

“My true being is unborn and changeless. I am the Lord who dwells in every creature.” (§4.6)

“Few see through the veil of maya. The world, deluded, does not know that I am without birth and changeless.” (§7.25)

“When you make your mind one-pointed through regular practice of meditation, you will find the supreme glory of the Lord.” (Gita §8.8)

“The immature do not look beyond physical appearances to see my true nature as the Lord of all creation.” (§9.11)

“I am the goal of life, the Lord and support of all, the inner witness, the abode of all.” (§9.18)

I am not a religious person but nevertheless I find it fascinating that the ancient Upanishads and the Gita, as reflected in these quotes from the translations by Prabhavananda (1975) and Easwaran (2007), describe my experience better than contemporary Western philosophers.

The Atman transcends logic, reason, and knowledge. It is a Something about which nothing can be said, and as such it is described apophatically, by way of denial and religious superlatives. And since it cannot be described, it cannot be ascribed any change and it is therefore said to be immutable and eternal. Finally, it is identified with the ultimate nature of reality, which may remind one of panpsychism or idealism.

The Atman is described in some quotes as an inner witness, the eye of the eye, and the thing through which one sees, which are descriptions that many people can relate to. Yet, the Upanishads and the Gita insist that few people behold that mysterious inner reality. How can that be?

Philosophers of mind look at each other with a good measure of disbelief. It sometimes seems as if deep in their hearts, dualists believe that Dennett knows he is wrong, and that he advocates his ideas as some kind of intellectual game designed to annoy them. But the peer response with which I have met suggests the possibility that Dennett really does not see the problem dualists contemplate.

Functionalists are off the mark too. For decades they have been trying to explain how mechanical processes ought to amount to full consciousness, but they have failed to explain how it is that their colleagues stubbornly fail to see what anyone with a decent capacity to reason can easily see.

There is a fascinating problem of other minds expressed by Q(meta). Dualists have been afraid to address it because they cannot conceive of consciousness without that mysterious ingredient, and they dare not seriously argue that other people are not really conscious.

But that is a false and unwarranted conclusion. To conceive of people who deny the Hard Intuition as philosophical zombies is as ridiculous as conceiving of the not-a-Nothing, which some people report, as a sort of radioactive glow. And the Hindu scriptures express the problem in terms of an introspective skill that anyone may in principle realize.

In my own case, I first came to realize that there is something unmechanical about my mind after several years of meditation. David Chalmers, too, recounts starting off as a materialist and gradually turning to mind-body dualism (Chalmers, 1997a, pp. xiv, 357).

According to Einstein there is something essential about the Now which is just outside the realm of science, and according to Wittgenstein, it is a Something about which nothing can be said, but Q(meta) itself is a question that may and should be investigated philosophically and scientifically.

References

BonJour, L. (2007). Epistemological problems of perception. Stanford encyclopedia of philosophy. Retrieved Dec 31, 2015, from <http://plato.stanford.edu/archives/spr2013/entries/perception-episprob/>

Carnap, R. (1963). The Philosophy of Rudolf Carnap. P. A. Schilpp (Ed.). Cambridge University Press.

Chalmers, D. J. (1995). Facing up to the problem of consciousness. Journal of consciousness studies, 2(3), 200-219. Retrieved Aug 21, 2015, from <http://cogprints.org/316/>

Chalmers, D. J. (1997a). The conscious mind: In search of a fundamental theory. Oxford Paperbacks.

Chalmers, D. J. (1997b). Moving forward on the problem of consciousness. *Journal of Consciousness studies*, 4(1), 3-46. Retrieved Oct 3, 2015, from <http://cogprints.org/317/1/moving.html>

Chomsky, N. (2009). The mysteries of nature: How deeply hidden?. *The Journal of Philosophy*, 167-200.

Chomsky, N., & Krauss, L. M. (2015) Chomsky & Krauss: An Origins Project Dialogue - (Part 2/2). [Video file] Retrieved from <https://www.youtube.com/watch?v=tbxp8ViBTu8&t=54s>

Dehaene, S. (2014). *Consciousness and the brain: Deciphering how the brain codes our thoughts*. Penguin.

Dennett, D. C. (1979). On the absence of phenomenology. In *Body, mind, and method* (pp. 93-113). Springer Netherlands.

Dennett, D. C. (1993). *Consciousness explained*. Penguin UK.

Dennett, D. C. (2001). The zombic hunch: Extinction of an intuition?. *Royal Institute of Philosophy Supplement*, 27-44. Retrieved Jan 30, 2016, from https://mitpress.mit.edu/sites/default/files/9780262541916_sch_0001_0.pdf

Dennett, D. C. (2003). The illusion of consciousness [Video file]. Retrieved from http://www.ted.com/talks/dan_dennett_on_our_consciousness

Descartes, R. (1637). *Discourse on method*. Retrieved Sep 27, 2015, from <http://www.gutenberg.org/files/59/59-h/59-h.htm>

Deutsch, D. (2004). It from qubit. *Science and ultimate reality: Quantum theory, cosmology, and complexity* (Festschrift for John A. Wheeler), 90-102.

Easwaran, E. (2007). *The Bhagavad Gita*. Nilgiri Press.

Einstein, A., Podolsky, B., & Rosen, N. (1935). Can quantum-mechanical description of physical reality be considered complete?. *Physical review*, 47(10), 777. Retrieved Jun 12, 2016, from <http://journals.aps.org/pr/pdf/10.1103/PhysRev.47.777>

Feynman, R. P. (1967). The character of physical law. MIT Press.

Feynman, R. P. (1982). Simulating physics with computers. *International Journal of Theoretical Physics*, 21(6), 467-488.

Hatfield, G. (2014). René Descartes. *Stanford Encyclopedia of Philosophy*. Retrieved Aug 26, 2015, from <http://plato.stanford.edu/archives/fall2015/entries/descartes/>

Hawking, S. (2010). 10 Questions for Stephen Hawking. *TIME*.

Horst, Steven. (2015). The Computational Theory of Mind. *Stanford Encyclopedia of Philosophy*. Retrieved Aug 13, 2015, from <http://plato.stanford.edu/archives/sum2015/entries/computational-mind/>

Humphrey, N. (1986). The Ghost in the Machine [Television series episode]. In *The inner eye*. Retrieved Jan 20, 2016, from https://www.youtube.com/watch?v=uZzv_R_51zw&t=47m30s

Janiak, A. (2014). Newton's philosophy. *Stanford Encyclopedia of Philosophy*. Retrieved Dec 25, 2015, from <http://plato.stanford.edu/archives/sum2014/entries/newton-philosophy/>

Jensen, K., & Karch, A. (2013). Holographic dual of an Einstein-Podolsky-Rosen pair has a wormhole. *Physical review letters*, 111(21), 211602.

Karpathy, A., & Fei-Fei, L. (2015). Deep visual-semantic alignments for generating image descriptions. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 3128-3137). Retrieved Aug 12, 2015, from <http://cs.stanford.edu/people/karpathy/deepimagesent/>

Kirk, R. (2015). Zombies. *Stanford encyclopedia of philosophy*. Retrieved June 2, 2017, from <https://plato.stanford.edu/archives/spr2015/entries/zombies/>

- Kuhn, R. L. (2011). What Is the Meaning of Consciousness? [Television series episode]. In Closer To Truth. Retrieved Aug 10, 2015, from <http://www.closetotruth.com/interviews/2731>
- Kurzweil, R. (2012). How to create a mind: The secret of human thought revealed. Penguin.
- Leibniz, G. W. (1898). The Monadology (R. Latta, Trans.). Retrieved Sep 18, 2015, from <http://home.datacomm.ch/kerquelen/monadology/printable.html>
- Lloyd, S. (2013). The universe as quantum computer. arXiv preprint arXiv:1312.4455.
- Moscow Center for Consciousness Studies (2012, June 21). Interview with Dennett on the Mind [Video file]. Retrieved from: https://www.youtube.com/watch?v=5_xlusVsA3Q&t=18m01s
- Moscow Center for Consciousness Studies (2013, July 1). David Chalmers's Theory of Consciousness [Video file]. Retrieved from: <https://www.youtube.com/watch?v=ijERcw0DMH8&t=1h21m>
- Oppy, G., & Dowe, D. (2011). The Turing Test. Stanford Encyclopedia of Philosophy. Retrieved Aug 10, 2015, from <http://plato.stanford.edu/archives/spr2011/entries/turing-test/>
- Prabhavananda, S., & Manchester, F. (1975). The Upanishads: Breath of the Eternal.
- Revonsuo, A. (1995). Consciousness, dreams, and virtual realities. *Philosophical Psychology*, 8(1), 35-58.
- Searle, J. R. (1984). *Minds, brains, and science*. Harvard University Press.
- Searle, J. R. (1997). *The mystery of consciousness*. New York Review of Books.
- Sytsma, J., & Machery, E. (2010). Two conceptions of subjective experience. *Philosophical Studies*, 151(2), 299-327. Retrieved Jan 21,

2016, from

http://philsci-archive.pitt.edu/4888/1/Two_Conceptions_of_Subjective_Experience.pdf

Turing, A. M. (1936). On computable numbers, with an application to the Entscheidungsproblem. *J. of Math*, 58(345-363), 5.

Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, 433-460. Retrieved Aug 12, 2015, from <http://cogprints.org/499/1/turing.HTML>

Yudkowsky, E. (2010). Eliezer Yudkowsky and Massimo Pigliucci on the Intelligence Explosion.

Wired UK (2015, June 22). DeepMind: Breakout [Video file]. Retrieved from <https://www.youtube.com/watch?v=Q70uIPJW3Gk>

Wittgenstein, L. (2009). *Philosophical Investigations*, revised 4th edition by PMS Hacker and Joachim Schulte. Wiley-Blackwell.