Max Velmans

Interview 16/01/15

Richard Bright: I would like to begin by asking, what is your definition of consciousness?

Max Velmans: I think that Tom Nagel's suggestion that consciousness is "what it's like to be something" is a useful place to start. But if we want to define what "consciousness" is in a more precise way we have to edge up to it more slowly. As with any term that refers to something that one can observe or experience it is useful, if possible, to begin with an ostensive definition, i.e. to "point to" or "pick out" the phenomena to which the term refers and, by implication, what is excluded.

Normally we point to some entity, event or process that we observe or experience. The term "consciousness" however refers to experience itself. Rather than being exemplified by a particular thing that we observe or experience, it is exemplified by all the things that we observe or experience. Something happens when we are conscious that does not happen when we are not conscious—and something happens when we are conscious of something that does not happen when we are not conscious of that thing. We know what it is like to be conscious when we are awake as opposed to not being conscious when in dreamless sleep. We also know what it is like to be conscious of something (when awake or dreaming) as opposed to not being conscious of that thing.

This everyday understanding of consciousness based on the presence or absence of experienced phenomena provides a simple place to start. A person, or other entity, is conscious if they experience *something*; conversely, if a person or entity experiences nothing they are not conscious. Elaborating slightly, we can say that when consciousness is present, phenomenal content (consciousness of something) is present. Conversely, when phenomenal content is absent, consciousness is absent.

That said, Eastern philosophies also refer to a state of "pure consciousness," without any phenomenal contents, although various characterisations are nevertheless offered of this state, such as *sat-chit-ananda* (being-consciousness-bliss) in Hindu thought, or *sunyata* (emptiness) in Mahayana Buddhism. Such states are of central importance to the practice of meditation, but, as these possibilities do not have a direct bearing on the problems of defining consciousness within the Western tradition, I will leave them to one side for now, without dismissing them.

Thinking of consciousness as *phenomenal consciousness* stays very close to everyday usage and it provides a simple place of departure on which widely divergent theories about consciousness can agree. For example, even eliminative-reductionists such as Dan Dennett agree that that conscious phenomenology *seems* to exist, and this provides the point of departure for his attempts at phenomenal elimination/reduction.

It also makes sense to stay as close as possible to everyday, natural language usage for related terms. In common usage, the term "consciousness" is often synonymous with "awareness", "conscious awareness", and “experience”. For example, It makes no difference in most contexts to claim that I am "conscious of" what I think, "aware of" what I think, "consciously aware" of what I think, or that I can “experience” what I think. Consequently, to minimise confusion, it is important not to load these terms with added meanings that are peculiar to a given theoretical position. This applies equally to the “contents of consciousness”. The "contents of consciousness" encompass all the phenomena that we are conscious of, aware of, or experience. These include not only experiences that we commonly associate with ourselves, such as thoughts, feelings, images, dreams, body sensations and so on, but also the experienced three-dimensional world (the phenomenal world) beyond the body surface.

However some terminological distinctions are important. In some older writings, for example in the work of Descartes, "consciousness" is not clearly differentiated from "mind." Given the extensive, current evidence for preconscious and unconscious mental processing, this usage is too broad. How phenomenal consciousness relates to preconscious and unconscious mental processing is now a major topic for psychological research. To avoid confusion, and to enable such research, it is important to reserve the term "mind" for psychological states and processes that may or may not be "conscious".

Descartes also famously believed thought to epitomise the nature of consciousness, and consequently defined it as a “substance that thinks” (*res cogitans*), which distinguishes it (in his view) from material substance that has extension in space (*res extensa*). Modern psychology accepts that verbal thoughts (in the form of phonemic imagery or ‘inner speech’) are amongst the contents of consciousness. However it does not accept that thoughts exemplify all conscious contents. Unlike thoughts, pains, tactile sensations, itches and other body experiences appear to have both spatial location and extension in different regions of the body, and the sights and sounds of the experienced external world (the phenomenal world) appear to have locations and extensions in a surrounding three-dimensional space. These interoceptive and exteroceptive experiences also differ widely from each other and many descriptive systems have been developed for investigating their phenomenology (in studies of visual and auditory perception, emotion, pain, and so on). It should be evident that such developments in phenomenology are an essential first step in characterising what it is about consciousness that needs to be explained—and that restricting the phenomenology of “consciousness” to the phenomenology of “thought” is too narrow.

In other, more modern writings, "consciousness" is sometimes taken to be synonymous with "self-consciousness". As one can be conscious of many things other than oneself (other people, the external world, etc.), this usage is also too narrow. To allow a clear distinction between consciousness of oneself and consciousness of things other than oneself, it makes more sense to reserve the term “self-consciousness” for a special form of reflexive consciousness in which the object of consciousness is the self or some aspect of the self.

The term "consciousness" is also commonly used to refer to a state of wakefulness. Being awake or asleep or in some other state such as coma clearly influences what one can be

conscious of. However these global states have a complex relationship to phenomenal consciousness. When sleeping, for example, one can still have visual and auditory experiences in the form of dreams. Conversely, when awake there are many things at any given moment that one does not experience. So in a variety of contexts it is necessary to distinguish "consciousness" in the sense of "phenomenal consciousness" from wakefulness and other states of arousal, such as dream sleep, deep sleep, and coma.

Finally, "consciousness" is sometimes used to mean "knowledge", in the sense that if one is conscious of something one also has knowledge of it. This is an important feature of consciousness. However, at any moment, much knowledge is unconscious, or implicit (for example, the knowledge gained over a lifetime, stored in long-term memory). So consciousness and knowledge cannot be co-extensive.

The above, broad definitions and distinctions have been quite widely accepted in the contemporary scientific literature—although by no means universally. Agreeing on definitions is important. Once a given reference for the term "consciousness" is fixed in its phenomenology, the investigation of its nature can begin, and this may in time transmute the meaning (or sense) of the term. As Dewey (1910) noted, to grasp the meaning of a thing, an event or situation is to see it in its relations to other things—to note how it operates or functions, what consequences follow from it, what causes it, and what uses it can be put to. Thus, to understand what consciousness is, we need to understand what causes it, what its function(s) may be, how it relates to nonconscious processing in the brain, and so on. As our scientific understanding of these matters deepens, our understanding of what consciousness is will also deepen. A similar transmutation of meaning (with growth of knowledge) occurs with basic terms in physics such as "energy", and "time."

RB: Is consciousness fundamental or incidental?

MV: This question can be interpreted in two ways. It could mean: (a) “Is consciousness fundamental or incidental to the functioning of the human mind”? Or (b) “Is consciousness a fundamental or incidental feature of nature?” As this last sense of the question relates closely to later questions that you ask about reflexive monism, panpsychism, continuity theory versus discontinuity theory, and the “hard problem” I will focus just on (a) here.

Whether consciousness is fundamental or incidental to the functioning of the human mind obviously depends on the function of consciousness itself—a matter that has been extensively investigated and debated within the consciousness studies field. For example, the suggestion that consciousness is an evolved function that is necessary to deal with any information that is complex or novel dates back to the naturalist George Romanes (a friend of Darwin) in 1885. And from the early 1960’s onwards, consciousness has been claimed by one or another psychological theory to be necessary to deal with novelty or complexity at every major stage of information processing (ranging from input, storage, and transformation to output) for example, in complex pattern recognition, learning, the operations of working memory, thinking, planning, problem solving, and complex, novel motor control such as speech production. If so, consciousness would play a central role in

the functioning of the human mind, making it, in this sense, fundamental to its operations. This view continues to be developed currently, for example in the work of Bernard Baars and Stanislas Dehaene on the functioning of the brain's "global workspace", which they identify with the functions of consciousness.

Unfortunately, doubts about such proposals have an equally distinguished history, dating back to the naturalist Thomas Huxley (another friend of Darwin) in 1874, based on his observations of the complex unconscious functioning of animals in which consciousness appeared to play no role. Rather, "The consciousness of brutes would appear to be related to the mechanism of their body simply as a collateral product of its working, and to be as completely without any power of modifying that working as the steam-whistle which accompanies the workings of a locomotive engine is without influence upon its machinery"—a position later known as epiphenomenalism. If so, consciousness would be incidental to the brain's functioning, although Huxley did accept that consciousness in humans "symbolise" molecular changes in their brains.

In 1958, the eminent neuroscientist Carl Lashley came to a similar conclusion about human mental functioning, as did the distinguished cognitive psychologist George Miller in 1962, based on an examination of his own experience. As he noted, "The fact that the process of thinking has no possible access into consciousness may seem surprising at first, but it can be verified quite simply. At this moment, as you are now reading, try to think of your mother's maiden name. What happened? What was your conscious awareness of the process that produced the name? Most persons report they had feelings of tension, of strain unrelated to the task, and then suddenly the answer was there in full consciousness. There may have been a fleeting image or two, but they were irrelevant. Consciousness gives no clue as to where the answer comes from; the processes that produce it are unconscious. It is the result of thinking, not the process of thinking, that appears spontaneously in consciousness." (Miller, 1962, p. 71) And, "What is true of thinking and of perceiving is true in general. We can state it as a general rule. No activity of mind is ever conscious. In particular, the mental processes involved in our desires and emotions are never conscious. Only the end product of these motivational processes can ever become known to us directly." (Ibid, p. 72)

I've carried out pretty extensive work on this issue myself. For example in my 1991 target article on "Is human information processing conscious?" (In the *Behavioral and Brain Sciences*) The view that consciousness must have a third-person causal role is supported by conventional evolutionary theory. After all, if it did not enhance inclusive fitness how could it have evolved? However, if one examines human information processing purely from a third-person perspective, consciousness does not seem to be necessary for any form of processing. As far as we know, the classical physical world causally closed. The operation of minds and brains seems to be explainable entirely in functional or physical terms that make no reference to what we experience. Once the processing within a system required to perform a given function is sufficiently well specified in procedural terms, one does not have to add an "inner conscious life" to make the system work. In principle, the same function, operating to the same specification, could be performed by a nonconscious machine.

Likewise, if one inspects the operation of the brain from the outside, no subjective experience can be observed at work. Nor does one need to appeal to the existence of subjective experience to account for the neural activity that one can observe. The neural correlates of consciousness already fill any "gaps" that might potentially be filled by consciousness in the activities of brain.

It turns out that the experimental and introspective evidence regarding how phenomenal consciousness actually relates to so-called "conscious processing" in humans deepens this puzzle. The detailed operations of most processes that we think of as "conscious" are not available to introspection. In stimulus identification and selection one is not aware of performing feature analysis, accessing long-term memory traces, or making assessments of the relative importance of preconscious stimuli. When remembering, one has no awareness of processes that perform memory search or retrieval. The phonemic images that constitute verbal thinking or "inner speech" give little information about the complex information transformations required to solve problems. And the detailed motor programmes controlling the musculature in speech or in complex adjustments to a changing environment have little manifestation in awareness. Rather, what enters awareness appears to *result* from such "conscious processing." The entities we perceive are the result of prior feature analysis and feature integration, and the names we assign to such entities "symbolise" the fact that these have been matched to long-term memory traces in a particular way. The events we remember *have been* searched for and retrieved (from long-term memory. And when we speak, the words that we hear ourselves utter are the result of *prior* semantic, syntactic and phonemic planning, and consequent motor control. In short, once one examines the timing of the experiences that accompany "conscious processing," the experiences seem to come *too late* to affect the processing to which they most obviously relate. The evidence for this is staring us in the face. For example, by the time you are conscious of this sentence (in the form of "inner speech") you will already have carried out the pattern recognition of individual words, and the semantic and syntactic analysis required to integrate the word meanings into a meaningful sentence.

When I examined the literature closely I also found that debates about the function of consciousness typically conflate three different senses in which a process can be said to be "conscious". A mental process might be conscious: (a) in the sense that one is conscious *of* the process (b) in the sense that the operation of the process *is accompanied by* consciousness (of its results) and (c) in the sense that consciousness enters into or *causally influences* the process.

What's this about? Ask yourself what's conscious about "conscious verbal thought". We do not have introspective access to how the preconscious cognitive processes that enable thinking result in conscious thoughts in the form of "inner speech." However, the content of such thoughts and the sequence in which they appear does give some insight into the way the cognitive processes (of which they are manifestations) operate over time in problem solving, thinking, planning and so on. Consequently such cognitive processes are partly conscious in sense (a), but only in so far as their detailed operation is made explicit in conscious thoughts, thereby becoming accessible to introspection and consequent report.

Many psychological processes are conscious in sense (b), but not in sense (a)—that is, we are not conscious of how the processes operate, but we are conscious of their results. This applies to perception in all sense modalities. When consciously reading this sentence for example you become aware of the printed text on the page, accompanied, perhaps, by inner speech (phonemic imagery) and a feeling of understanding (or not), but you have no introspective access to the processes which enable you to read. Nor does one have introspective access to the details of most other forms of cognitive functioning, for example to the detailed operations which enable "conscious" learning, remembering, engaging in conversations with others and so on.

Crucially, having an experience that gives some introspective access to a given process, or having the results of that process manifest in an experience, says nothing about whether that experience carries out or controls that process. That is, whether a process is "conscious" in sense (a) or (b) needs to distinguished from whether it is conscious in sense (c). Indeed, it is not easy to envisage how the experience that makes a process conscious in sense (a) or (b), *could* make it conscious in sense (c). Consciousness of a physical process does not make consciousness responsible for the operation of that process (watching paint dry does not actually make it dry on the wall). So, how could consciousness of a mental process carry out the functions of that process? Alternatively, if conscious experience results from a mental process it arrives *too late* to carry out the functions of that process.

Such considerations leave one with what I have referred to as a *causal paradox*. That brain states have a causal influence on conscious experiences seems undeniable. As Thomas Huxley pointed out in 1874, one has only to stick a pin in oneself to give a sufficient demonstration. But if consciousness is viewed in traditional dualist terms, how brain states cause conscious experiences seems inexplicable. Neural causes might have neural and other physical effects, but how could something "objective" and "physical" produce a "subjective experience"?

Nor is it clear how consciousness might influence processing in the brain. We normally take it for granted that we have a conscious mind that controls our voluntary actions and this view is fundamental to our ethics, politics and legal systems. But how the conscious mind exercises its influence is not easy to understand. Viewed from a first-person perspective, consciousness appears to be necessary for most forms of complex or novel processing. But, viewed from a third-person perspective, consciousness does not appear to be necessary for any form of processing, as there are no "gaps" in the chain of neurophysiological events that require the intervention of consciousness to make the brain work.

So—to get back to your question, if the mind is viewed from a first person perspective, consciousness appears to be fundamental, while viewed from a third-person perspective, it appears to be incidental (all one can observe are the operations of brain). I think that there is a way to resolve this paradox in an intuitively plausible way that avoids both epiphenomenalism and materialist reductionism, but this requires us to shift our explanatory framework to one in which first-person and third-person views of the mind's operations are treated as *complementary and mutually irreducible*. On this view, consciousness is central to human mental functioning viewed from a first-person

perspective, while from a third-person perspective, mental functioning can be explained purely by the operations of brain. However, a *complete* description of the nature of mind requires both perspectives (that's why they are complementary). This has many interesting consequences, but takes quite a lot of explanation—so, as you have not asked me about this aspect of my work, I won't go into it here. Readers can however have look at my 2002 Journal of Consciousness Studies target article on "How could conscious experiences affect brains?" or chapters 10 and 13 of my book *Understanding Consciousness: Second Edition* (2009) to see where all this goes.

RB: In the light of recent discoveries in neuroscience, does the concept of consciousness need revising?

MV: No and Yes. As I've suggested in my definition of consciousness above, thinking of consciousness as *phenomenal consciousness* 'points to' or 'picks out' the phenomena of central interest in Western discussions and debates, and that won't change as neuroscience advances. However, as our scientific understanding of the way that brain states influence conscious experience deepens, our understanding of what "consciousness" is will also necessarily deepen, as the relation of consciousness to the physical world is part of the connotative meaning or 'sense' of the term. And there are many other things to understand about "consciousness" that will affect our concept of it, for example, its functions, its apparent causal interactions with the brain, the relation of conscious to unconscious processing, and so on.

It also needs to be said that such expanded forms of understanding are bi-directional. The principle that first- and third person perspectives on the mind are complementary and mutually irreducible also allows that a more precise and inclusive first-person exploration of conscious phenomenology can encourage a deeper understanding of its neural accompaniments, leading to a form of "experiential neuroscience"—to which we return below.

I want to stress, however, that neuroscience will *never* demonstrate consciousness to be nothing more than a state or function of the brain, for the simple reason that *causation*, *correlation* and *ontological identity* are very different relationships. Such a reduction would require one to establish ontological identity between conscious experiences and their associated causes and/or correlates. However, third-person investigations of the brain can only discover what those neural causes and correlates are. From the mid 20th Century there have extensive efforts to get around this theoretically, which I have reviewed in depth in chapters 3, 4 and 5 of *Understanding Consciousness*, so I won't go into it here. Readers can also look at *How to define consciousness—and how not to define consciousness*, or my review of Stan Dehaene's (2014) book on *Consciousness and the Brain*, in which I give many reasons for separating the fascinating advances in the neuroscience of consciousness over the last 25 years from an unjustifiable, materialist-reductionist philosophy.

RB: The standard reductionist approach to consciousness portrays it as a state or function of the brain, that is, sensory inputs to the brain are processed to the point where they become

a conscious experience in the brain. You have developed a different theory of consciousness called reflexive monism. What is reflexive monism?

MV: Monism is the view that the universe, at the deepest level of analysis, is one thing, or composed of one fundamental kind of stuff. This is usually contrasted with Substance Dualism, the view found, for example in the writings of Plato and Descartes that, fundamentally, the universe is composed of two kinds of stuff, physical stuff and the stuff of soul, mind or consciousness. Reflexive Monism (RM), a philosophical position I developed in *Understanding Consciousness*, 2000, 2009 is a modern version of an ancient view that the basic stuff of which the universe is composed has the potential to manifest both physically and as conscious experience (a dual aspect monism in the tradition of Spinoza). In its evolution from some primal undifferentiated state, the universe differentiates into distinguishable physical entities, at least some of which have the potential for conscious experience, such as human beings. While remaining embedded within and dependent on the surrounding universe and composed of the same fundamental stuff, each human, equipped with perceptual and cognitive systems has an individual perspective on, or *view* of, both the rest of the universe and him or her self. In this sense, each human participates in a process whereby the universe differentiates into parts and becomes conscious in manifold ways of itself, making the entire process reflexive. A detailed analysis of Reflexive Monism is given in my book *Understanding Consciousness Edition 2* (2009). Central features of this theory are also summarised in "Reflexive monism", JCS, 2008, and developed in ways not entirely covered by my book in, "Reflexive monism: psychophysical relations among mind, matter and consciousness", JCS, 2012. Its similarities and differences to Eastern philosophy, particularly Advaita Vedanta are discussed in detail in "How to arrive at an Eastern place from a Western direction" (2013). And they are also discussed in an online (2014) talk on "From West towards East in five simple steps" available here.

The paragraph above gives a rough, global description of RM, but the system can really only be properly understood if one examines how it deals with the many problems of consciousness and how they relate to each other in detail. For that, one would have to read the book. Given what we've already discussed above, it's nevertheless worth saying a little more about its two central features--monism and reflexivity.

My own journey into RM started in 1975 with a simple observation about perception and my own conscious experience. Before I started serious work on consciousness my own pre-theoretical assumptions about what and where it is in relation to the brain and physical world were entirely conventional. For example when I looked at this print on the page, I took it for granted that this print was physical, 'out-here' in the physical world, and that, in perception, light reflectances from this page were picked up by my visual system 'over-here' and subsequently processed in my brain. I also took it for granted that my conscious experiences *of* this print were 'over-here' somewhere in my brain (or maybe nowhere, if I adopted dualism). When I thought about it again though, it struck me that these assumptions about my conscious experience didn't correspond in any way to its *phenomenology*! For example, when I look at this page, the only visual experience *of* print that I have is the print as-seen out-here on this page—and I don't have any *added* experience of print 'nowhere' or in my brain. And that applies to the whole of exteroception, i.e. to the entire external, three-dimensional phenomenal world. Pains,

tactile and other body sensations are similar. For example if I stab my finger with a pin, I feel the pain out-there in my finger, not in my brain. Indeed, experiences that seem to be in my head or brain or maybe 'nowhere' are relatively few—although there are some—thoughts for example, and certain forms of visual imagery.

This suggests that normal perception is *reflexive*. Physical stimuli out-there on the page are experienced as being out-there, stimuli originating in my body are experienced as being located in the body, and events originating in the brain itself such as the cognitive processing that results in verbal thoughts are experienced as being in the head or brain. Crucially, in terms of *phenomenology*, there is no difference between "experiences *of* events" and (the same) "events as-experienced". It follows that the phenomenal world that, in everyday life, we think of as the "physical world" is actually *part of* conscious experience. It was never *apart from* it—although the phenomenal physical world is not, of course, the same as the world described by physics (in terms of quantum mechanics, relativity theory and so on).

Many surprising consequences follow from this, outlined in an initial way, for example, in *Consciousness, brain and the physical world* (1990)--my first publication on this subject. I can't elaborate on these consequences here, but it should be evident that this simple, reflexive way of viewing perception provides a microcosmic example of how reflexivity might operate within reflexive monism.

The *monism* in reflexive monism derives, in the first instance, from a similarly simple source—in this case from a simple way to make sense of what can be observed or plausibly inferred about the workings of the human mind.

It is widely assumed for example that for each, normal conscious experience one can expect to find a distinct physical correlate. Although we do not have complete knowledge about the physical nature of these correlates, there are four plausible, functional constraints imposed by the phenomenology of consciousness itself:

1. The representational constraint. Normal human conscious experiences are representational (phenomenal consciousness is always *of* something). Given this, it is plausible to assume that the physical correlates of such experiences are representational states.

2. The identical referent constraint. A representational state must represent *something*. For a given physical state to be the correlate of a given experience it is plausible to assume that it represents the *same* thing (otherwise it would not be the correlate of *that* experience).

3. The information preservation constraint. For a physical state to be the correlate of a given experience, it is reasonable to suppose that it has the same "grain". That is, for every discriminable attribute of experience there will be a distinct, correlated, physical state. As each experience and its physical correlate represent the same thing it follows that each experience and its physical correlate encodes the same information about that thing. That is, they are representations with the same *information structure*.

4. Orderly mapping. It is reasonable to assume that the formatting of neurally encoded information relates to the formatting of corresponding, phenomenally encoded information in an orderly way, with discoverable neural state space/phenomenal space mappings. An obvious example would be the way that information about spatial location and extension encoded in the brain is mapped into the 3D phenomenal space that we ordinarily experience.

Ever since the pioneering work of Gustaf Fechner (1860), these assumptions have largely been taken for granted in psychological theory. The assumption that experiences and their physical correlates encode identical information also marks an important point of convergence between otherwise divergent theories about the nature of consciousness. This assumption is implicit, for example, in eliminativist and reductionist theories of consciousness. So—up to this point, my analysis is entirely conventional.

However, accepting that conscious experiences and their physical correlates encode identical information doesn't alter the fact that, viewed from the perspective of those who embody them, the operations of mind take the form of conscious experiences—while the operations of mind viewed from the outside just seem to be the operations of brains (or some physical aspect of brains).

If we can't eliminate or ignore the facts viewed from either perspective (as I believe), then
A straightforward way to make sense of this situation is to assume that there is one underlying mental process that grounds and connects these two ways of knowing it—a position that can be formally described as *ontological monism combined with epistemological dualism.*

Taken together, these points also suggest that mind can be thought of as a form of information processing, and the information displayed in experiences and their physical correlates can be thought of as two manifestations of this information processing—which makes this a *dual-aspect theory of information processing*.

However, this does not fully specify the ontology of the mind. Information processing needs to be encoded in some medium that is capable of carrying out that processing. Given this, what kind of medium is the mind? If first- and third-person perspectives (on the mind) are complementary and mutually irreducible, then the nature of the mind is revealed as much by how it appears from one perspective as the other. If so, the nature of mind is not *either* physical *or* conscious experience; it is at once physical *and* conscious experience. For lack of a better term we may describe this nature as *psychophysical.* If we combine this with the features above, we can say that mind is a psychophysical process that encodes information, developing over time.

If one thinks of the human mind as an entirely natural manifestation of what nature is like (as I do), this also opens up the possibility that the fundamental stuff of the universe is itself psychophysical (as Gustaf Fechner supposed), but we will return to that in your questions about panpsychism and continuity theory below.

RB: Is there a 'Hard Problem'?

MV: It has been recognised that there are hard problems surrounding consciousness ever since Descartes suggested that consciousness (*res cogitans*) and the material world (*res extensa*) causally interact. For example in her correspondence with Descartes, Princess Elizabeth in 1643 already expressed her doubts that something "extended" could ever interact with something that "thinks". Stated in modern terms, how could *electrochemistry* give rise to *subjective experiences*? Conversely, how could experienced *wishes* or *desires* affect the behaviour of *neurons*? Given this, it's not surprising that Spinoza (1677) and Leibniz (1696) judged the causal interaction of *res cogitans*, and *res extensa* to be literally inconceivable. In my review of the field in my 1991 *BBS* target article I gave many additional reasons to be doubtful about (a) such causal interactions, and (b) the reducibility of phenomenal consciousness to states or functions of the brain (see above).

Chalmers' (1995) "easy" versus "hard" problem distinction in the *Journal of Consciousness Studies* nevertheless provided a useful reminder that a purely third-person functional analysis of human information processing cannot reveal what it is like to have a subjective experience or explain why it arises. However, his division of the problems of consciousness into the "easy" problems and the "hard" one was, in turn, an oversimplification. As Chalmers himself accepted, even so-called "easy" (empirically researchable) problems can in practice be very difficult to solve. It may also be that the "hard" problem only seems unusually hard because we have been thinking about it in the wrong way. If so, changing some of our unexamined assumptions might be all we need to make the problem "easy".

In my own work, for example in *Understanding Consciousness* 2009, I note that Western science usually takes the existence of matter for granted, while the existence of consciousness is regarded as mysterious. Consequently, the conventional "hard problem" refers to the difficulty of understanding how consciousness arises from insentient physical matter, or, in other versions, about the seeming irreducibility of first-person accounts of conscious experience to third-person descriptions of the brain. But in truth, the existence of matter is as mysterious as the existence of consciousness, and there are similarly hard problems in physics. Why, for example, should electricity flowing down a wire be accompanied by a magnetic field around the wire, why should electrons sometimes behave as waves and at other times as particles, and why there should be any matter in the universe at all?

We simply assume these to be natural facts that we can observe in the world. We can try to explain them by incorporating them into some body of theory, but we do not usually agonize over their *existence*. If it turns out that first-person and third-person accounts of the mind, along with the aspects of mind that they describe are *complementary and mutually irreducible* (see above) one would not expect to be able to derive one aspect from, or reduce one aspect to the other. It might just be a natural fact about the world that certain forms of brain functioning are accompanied by certain forms of first-person experience. That would require us to change a few of our pre-theoretical assumptions about the nature of matter and its relationship to consciousness, and we would still have to investigate the principles that govern the consciousness-brain relationship in great detail. But the fact that

given conscious states accompany certain forms of brain functioning would then be “hard” to understand in the same sense as many facts in physics.

While the parallels are not exact, wave-particle complementarity in quantum mechanics provides a rough analogy. One can relate wave and particle properties of electrons to each other with great precision, but within physics, neither is regarded as more basic than, reducible to, or supervenient on the other. Rather, such properties are regarded as *complementary and mutually irreducible*, just as first- and third-person observable aspects of mind are treated as complementary and mutually irreducible in my account of their causal interactions (above). And physics has to grapple with the very same issue of how to specify what it is that these complementary properties *are properties of.* Physics typically opts for descriptions that somehow combine wave and particle-like aspects, for example, describing electrons as “wave packets” or “electron clouds”. In similar fashion, in my own work, I opt to describe the fundamental nature of mind as “psychophysical”.

Without foreclosing on the possibility of a deeper understanding of electrons, e.g. in a mathematical form, quantum mechanics accepts that there is something deeply mysterious about the fundamental nature of matter. Without foreclosing on the possibility of a deeper understanding of mind, for example in terms of mathematically described bridging laws between conscious experiences and their neural accompaniments, I similarly accept that there is something deeply mysterious about the way that consciousness and the material forms with which they correlate arise from some “psychophysical” ground.

Given all this, rather than speaking of a single “hard problem”, it seems more useful to sort the problems of consciousness into those that require empirical advance, those that require theoretical advance, those that require a re-examination of some of our pre-theoretical assumptions, and those that require some combination of all three.

RB: Recent debates on consciousness has once again brought panpsychism into discussion. What is your view on panpsychism?

MV: My own work focuses primarily on human consciousness so none of it *requires* panpsychism to be true. RM would be internally consistent even if it applied only to human beings. My reason for developing it this way is that I did not want to confuse the problems surrounding human consciousness with the added difficulties of ascertaining its presence in nonhuman beings. That said, RM naturally accommodates panpsychism, and a combination of RM with panpsychism provides the most intellectually satisfying version of the theory. For example, if the universe, at the deepest level of analysis, is one thing, or composed of one fundamental *psychophysical* kind of stuff, the “hard problem” simply disappears, for the reason that both matter and its accompanying experience can then be seen as natural manifestations of that stuff (just as electricity and magnetism are natural manifestations of electromagnetism)—demonstrating that all one needs to make the “hard” problem “easy” is to change one of our pre-theoretical assumptions about what is basic. Rather than being a freak accident of nature, conscious human beings can then be seen as one of its infinitely varied, natural manifestations.

Of course, theoretical simplicity alone isn't enough to settle this question. But one's answer to your *next* question takes us closer.

RB: Related to the above question, can you say something about the distinction between continuity and discontinuity theories?

MV: I've written about the evolution of consciousness fairly extensively for example in the last chapter of *Understanding Consciousness* (both editions) and, most recently, in "The evolution of consciousness" (2012). In my surveys of the many attempts to make sense of the evolution of consciousness I observed that theories about the distribution of consciousness divide into *continuity* and *discontinuity* theories. Discontinuity theories all claim that consciousness emerged at a particular point in the evolution of the universe. They merely disagree about which point. Consequently, discontinuity theories all face the same problem. What switched the lights on? What is it about matter, at a particular stage of evolution, which suddenly gave it consciousness? Most try to define the point of transition in functional terms, although they disagree about the nature of the critical function. Some think consciousness "switched on" only in humans, for example once they acquired language or a theory of mind. Some believe that consciousness emerged once brains reached a critical size or complexity. Others believe it co-emerged with the ability to learn, or to respond in an adaptive way to the environment.

I argue in my work that such theories confuse the conditions for the *existence* of consciousness with the added conditions that determine the many *forms* that it can take. Who can doubt that verbal thoughts require language, or that full human self-consciousness requires a theory of mind? Without internal representations of the world, how could consciousness be *of* anything? And without motility and the ability to approach or avoid, what point would there be to rudimentary pleasure or pain? However, none of these theories explains what it is about such biological functions that suddenly switch on consciousness.

Continuity theorists do not face this problem for the simple reason that they do not believe that consciousness suddenly emerged at *any* stage of evolution. Rather, as Sherrington proposed in 1942, consciousness is a "development of mind from unrecognisable into recognisable." On this *panpsychist* or *panexperientialist* view, all forms of matter have an associated form of consciousness. In the cosmic explosion that gave birth to the universe, consciousness co-emerged with matter and co-evolves with it. As matter became more differentiated and developed in complexity, consciousness became correspondingly differentiated and complex. The emergence of carbon-based life forms developed into creatures with sensory systems that had associated sensory "qualia." The development of *representation* was accompanied by the development of consciousness that is *of* something. The development of *self-representation* was accompanied by the dawn of differentiated self-consciousness and so on. On this view, evolutionary theory can in principle account for the different *forms* that consciousness takes. But, consciousness, in some primal form, did not emerge at any particular stage of evolution. Rather, it was there from the beginning. Its emergence, with the birth of the universe is neither more nor less mysterious than the emergence of matter and energy.

Most discontinuity theorists take it for granted that consciousness could only have appeared (out of nothing) through some random mutation in complex life forms that happened to confer a reproductive advantage that can be specified in third-person functional terms. This deeply ingrained, pre-theoretical assumption has set the agenda for what discontinuity theorists believe they need to explain. Within cognitive psychology, for example, consciousness has been thought by one or another theorist to be necessary for every major phase of human information processing, for example in the analysis of complex or novel input, learning, memory, problem solving, planning, creativity, and the control and monitoring of complex, adaptive response. It should be apparent that continuity theory shifts this agenda. The persistence of different, emergent biological forms may be governed by reproductive advantage. If each of these biological forms has a unique, associated consciousness, then matter and consciousness co-evolve. However, conventional evolutionary theory does not claim that *matter itself* came into being, or persists through random mutation and reproductive advantage. According to continuity theory, neither does consciousness.

Which view is correct? One must choose for oneself. However, in the absence of anything other than arbitrary criteria for when consciousness suddenly emerged, continuity theory is arguably more elegant. Continuity in the evolution of consciousness favours continuity in the distribution of consciousness, although there may be critical transition points in the *forms* of consciousness associated with the development of life, representation, self-representation, and so on.

RB: William James proposed that the mind should be studied not only by way of behaviour and brain functions, but should include introspection. Scientists have taken up the first two very keenly, but have been reluctant to take up the method of introspection. Recently, the practice of what is called 'contemplative science' has been called for, a coming together of 'contemplative' (meditative) and 'scientific' methods of inquiry. How might science and contemplative practice collaborate in the study of consciousness and what, if any, do you see as the benefits of such a collaboration?

MV: Recent cultural acceptance of practices such as mindfulness training for enhancing normal mental functioning or ameliorating conditions such as depression and anxiety have attracted wide interest, for the reason that these are measurable behavioural changes. This has also led to a resurgence of interest in charting changes in brain states and functions associated with such practices—although studies of the effects of Eastern practices on brain and behaviour have actually existed over many decades. It's important to note too, that although these contemplative practices have their own distinctive characteristics and deserve to be studied in their own right, they are just one of many forms of first-person methods for exploring the nature of mind. Although introspection as such has been treated with suspicion over the last 100 years or so, neuropsychology has always accepted that one needs reports of first-person experience of a simple kind to provide a reference for neuroscientific investigations. One can't for example study the visual system without asking people what they see or getting them to respond to what they see in various experimental or clinical situations. More detailed examination of how conscious experiences of more

complex kinds relates to various aspects of mental information processing also extends over decades, e.g. to Ericsson and Simon's protocol analysis in 1980. In recent years there has been a resurgence of such interest, for example arising from Varela's "neurophenomenology" perhaps best exemplified in the work of Claire Petitmengin. The most systematic work I know of however is the "experiential neuroscience" developed by Donald Price and James Barrell, summarised in their 2012 book on the subject. This develops the principle that first- and third-person views of the mind are complementary, mutually irreducible and mutually informing (in the ways outlined above) in great detail.

RB: What are currently the most important questions, problems, or challenges confronting the understanding of consciousness, and what are the prospects for progress?

MV: Conventional scientific studies of consciousness within cognitive psychology and neuropsychology are ticking along nicely, and are starting to provide a deeper understanding of the "neural signatures" of consciousness within the normally functioning human brain, summarised for example in Dehaene's (2014) Consciousness and the Brain. Within conventional, materialist-reductionist culture however, we are still restricting our own understanding of consciousness both philosophically and scientifically, to a dehumanising and over-restricted view of what consciousness is and of how to explore it and transform it. Following Dan Dennett, for example, some scientists are still trying to persuade us that phenomenal consciousness is just an illusion, which, in my view, doesn't make any sense (see for example, my review of Humphreys 2014 book Soul Dust). This also discourages the development of first-person methods of the kind discussed in the question above. The challenge is to move to a more inclusive view of consciousness that is more in tune with our natural intuitions and more likely to encourage human flourishing, while remaining rigorous and consistent with the findings of science. This is likely to have broad, cultural consequences as well as philosophical/scientific ones that are much needed at this time.

RB: Finally, what are you working on now?

MV: Over the years I've been interested in doing what I can to foster consciousness studies (which didn't exist as a field of study when I first started working on it formally in 1975) while personally working towards a more integrated, inclusive understanding of it, for example through the *reflexive monism* that I developed. That work continues. For example, I am currently co-editing Edition 2 of the *Blackwell Companion to Consciousness*, which surveys the current state of the field, and I've also been commissioned to edit a four-volume work on *Consciousness* for Routledge's *Major Works* Series. My personal experiential and theoretical work is ongoing and continues to take me in interesting and sometimes surprising directions, for example towards the East … from a Westerly direction. It's a never-ending, fascinating journey!